

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 15:29:50 ; Search time 151 seconds
(without alignments)
23.757 Million cell updates/sec

Title: SEQ33

Perfect score: 72

Sequence: 1 cgywtiwcg 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 35872929 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	72	100.0	10	5	ABP53932 VEGFR-3 b
2	54	75.0	10	5	ABP53931 VEGFR-3 b
3	50	69.4	10	5	ABP53968 VEGFR-3 b
4	47.5	66.0	304	7	ADC87481 Human GPC
5	47	65.3	25	7	ADC99638 Cancer-re
6	46	63.9	17	6	Aao26074 Fc region
7	46	63.9	17	8	ADJ50741 Human ser
8	46	63.9	136	8	ADM87650 Human EST
9	45	62.5	20	5	Aau90545 Insulin/I
10	45	62.5	129	2	Aay59880 Human nor
11	45	62.5	828	6	ABU49938 Protein e
12	44.5	61.8	1024	5	ABBo4861 LDL recep
13	44	61.1	149	8	ADP29850 Human sec
14	44	61.1	166	6	AAE31487 Human but
15	44	61.1	115	5	AAU91293 Human NOV
16	43	59.7	13	6	Aao26093 Fc region
17	43	59.7	13	8	ADJ50760 Human ser
18	43	59.7	454	5	ABP73979 Candida a
19	43	59.7	474	6	ABU30004 Protein e
20	43	59.7	492	7	ADC97318 E. faecul
21	43	59.7	1084	8	ADL1869 P. aerugi
22	42.5	59.0	152	3	AAY94989 Human sec
23	42.5	59.0	152	5	ABb90081 Human pol
24	42.5	59.0	152	5	Aao17173 Human sec
25	42.5	59.0	152	5	ABG64784 Human alb

26	42.5	59.0	152	8	ADL78051 Albumin f
27	42.5	59.0	159	2	AAR66278 Therapeut
28	42.5	59.0	159	3	ABb12156 Hydrophob
29	42.5	59.0	159	4	AAW78581 Human pro
30	42.5	59.0	159	4	AAW73100 Human ang
31	42.5	59.0	159	5	AAO17198 Human sec
32	42.5	59.0	159	5	ABG64785 Human alb
33	42.5	59.0	159	8	ADL78052 Albumin f
34	42.5	59.0	159	8	ADN05188 Antipsori
35	42.5	59.0	161	4	AAW25822 Human pro
36	42.5	59.0	161	4	ABB12006 Human gli
37	42.5	59.0	161	4	AAW79565 Human pro
38	42.5	59.0	176	6	ABO07116 Novel hum
39	42.5	59.0	1024	5	ABBo4863 LDL recep
40	42.5	59.0	1765	2	AAV16572 Type 5 so
41	42.5	59.0	1765	2	AAV41668 Rat senso
42	42.5	59.0	1765	2	AAV06596 Rat sodiu
43	42.5	59.0	1765	4	AAAB20122 Rat sodiu
44	42.5	59.0	1765	4	AAAB20123 Rat sodiu
45	42.5	59.0	1765	7	ADD32192 Rat Na v

ALIGNMENTS

RESULT 1

ABP53932

ID ABP53932 standard; peptide; 10 AA.

XX AC ABP53932;

XX AC

DT 09-JAN-2003 (first entry)

XX AC

DE VEGFR-3 binding peptide SEQ ID NO:35.

XX AC

KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
cytostatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
vulnary; cell surface receptor; cancer; neovascularisation;
liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
diabetes; PDGF; platelet derived growth factor.
XX Homo sapiens.
OS Synthetic.
XX WO200257299-A2.
XX PD 25-JUL-2002.

XX 16-JAN-2002; 2002WO-IB0000099.

XX 17-JAN-2001; 2001US-0262476P.

XX (LUDW-) LUDWIG INST CANCER RES.

XX (LICN) LICENTIA LTD.

XX Alitalo K, Koivunen E, Kubo H;

XX WPI; 2002-691521/74.

XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
such as cancer and diseases of neovascularization.
XX Claim 13; Page 80; 149pp; English.
XX The present invention describes an isolated peptide (I) that binds to and
inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
antidiabetic and vulnary activities, and can be used in gene therapy.
CC Compositions and methods from the present invention are useful for
diagnosing, evaluating and treating disorders mediated by the activity of
the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 XX
 SQ Sequence 10 AA;

Query Match 100.0%; Score 72; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGYWLTIWGC 10
 |||||
 Db 1 CGYWLTIWGC 10

RESULT 2
 ABP53931
 ID ABP53931 standard; peptide; 10 AA.
 XX
 AC ABP53931;
 XX
 DT 09-JAN-2003 (first entry)
 XX
 DE VEGFR-3 binding peptide SEQ ID NO:34.
 XX
 KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
 KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
 KW cytotatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
 KW vulnary; cell surface receptor; cancer; neovascularisation;
 KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
 KW diabetes; PDGF; platelet derived growth factor.
 XX
 OS Homo sapiens.
 OS Synthetic.

XX
 FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "any amino acid"
 FT
 FT Misc-difference 10
 FT
 FT Misc-difference 10 /note= "any amino acid"
 FT

XX WO200257299-A2.
 XX
 XX 25-JUL-2002.
 XX
 PF 16-JAN-2002; 2002WO-IB0000099.
 XX
 PR 17-JAN-2001; 2001US-0262476P.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA (LICN) LICENTIA LTD.
 XX
 XX Alitalo K, Koivunen E, Kubo H;
 PI
 XX WPI; 2002-691521/74.
 DR

XX
 PT New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
 PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
 PT such as cancer and diseases of neovascularization.
 XX
 XX Claim 12; Page 80; 149pp; English.

XX
 PS The present invention describes an isolated peptide (I) that binds to and
 CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnary activities, and can be used in gene therapy.
 CC Compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 XX
 SQ Sequence 10 AA;

Query Match 75.0%; Score 54; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.67;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GYWLTIWG 9
 |||||
 Db 2 GYWLTIWG 9

RESULT 3
 ABP53968
 ID ABP53968 standard; peptide; 10 AA.
 XX
 AC ABP53968;
 XX
 DT 09-JAN-2003 (first entry)
 XX
 DE VEGFR-3 binding peptide SEQ ID NO:73.
 XX
 KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
 KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
 KW cytotatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
 KW vulnary; cell surface receptor; cancer; neovascularisation;
 KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
 KW diabetes; PDGF; platelet derived growth factor.
 XX
 OS Homo sapiens.
 OS Synthetic.

XX
 FH Key Location/Qualifiers
 FT Misc-difference 5.7 /note= "X is any amino acid"
 FT
 FT Misc-difference 9 /note= "X is any amino acid"
 FT

XX WO200257299-A2.
 XX
 XX 25-JUL-2002.
 XX
 PF 16-JAN-2002; 2002WO-IB0000099.
 XX
 PR 17-JAN-2001; 2001US-0262476P.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA (LICN) LICENTIA LTD.
 XX
 XX Alitalo K, Koivunen E, Kubo H;
 PI
 XX WPI; 2002-691521/74.
 DR

XX
 PT New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
 PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
 PT such as cancer and diseases of neovascularization.
 XX
 XX Disclosure; Page 147; 149pp; English.

XX
 PS The present invention describes an isolated peptide (I) that binds to and
 CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnary activities, and can be used in gene therapy.
 CC Compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a VEGFR-3 binding peptide, which is given in the
 CC exemplification of the present invention
 XX
 SQ Sequence 10 AA;

Query Match 69.4%; Score 50; DB 5; Length 10;
 Best Local Similarity 60.0%; Pred. No. 2.3;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGYWLTIWGC 10
 |||||
 Db 1 CGYWXWXXWC 10

RESULT 4
 ADC87481
 ID ADC87481 standard; protein; 304 AA.
 XX
 AC ADC87481;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Human GPCR protein SEQ ID NO:1934.
 XX
 KW human; GPCR; guanosine triphosphate-binding protein coupled receptor;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1270724-A2.
 XX
 PD 02-JAN-2003.
 XX
 PF 18-JUN-2002; 2002EP-00013517.
 XX
 PR 18-JUN-2001; 2001JP-00246789.
 XX
 PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.

PI Suwa M, Asai K, Akiyama Y, Aburatani H;
 XX WPI; 2003-315783/31.
 XX DR N-PSDB; ADC87480.
 XX
 PT New polynucleotide, useful for preparing a composition for treating a
 PT patient in need of increased or suppressed activity or expression of the
 PT guanosine triphosphate-binding protein coupled receptor.
 XX
 PS Claim 2; SEQ ID NO 1934; 28pp; English.
 XX
 CC The invention relates to a novel polynucleotide encoding a guanosine
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
 CC the invention may have a use in gene therapy. The polynucleotide and
 CC polypeptide are useful for preparing a composition for treating a patient
 CC in need of increased or suppressed activity or expression of the
 CC guanosine triphosphate-binding protein coupled receptor. The protein
 CC sequences shown in ADC85549-ADC87617 represent GPCR's of the invention.
 XX
 SQ Sequence 304 AA;

Query Match 66.0%; Score 47.5; DB 7; Length 304;
 Best Local Similarity 60.0%; Pred. No. 91;
 Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 CGYWLTIWGC 10
 |||||
 Db 108 CGFW-AVWGC 116

RESULT 5
 ADC99638
 ID ADC99638 standard; peptide; 25 AA.
 XX
 AC ADC99638;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Cancer-related Tie-1-binder peptide - SEQ ID 476.
 XX
 KW cytostatic; cancer; gene therapy; DGI-2; DGI-5; DGI-7; DGI-9; Hras;
 KW lepin; VEGF; vascular endothelial growth factor receptor; VEGF-R1;
 KW VEGF-R2; VEGF-R3; FLT1; FMS-related tyrosine kinase 1; FLK1; KDR;
 KW kinase insert domain protein receptor; EGFR; epidermal growth factor;
 KW FGFR1; fibroblast growth factor; tie-1.
 XX
 OS Unidentified.
 XX
 PN WO2003035839-A2.
 XX
 PD 01-MAY-2003.
 XX
 PF 24-OCT-2002; 2002WO-US034021.
 XX
 PR 24-OCT-2001; 2001US-0345471P.
 XX
 PA (DGI-) DGI BIOTECHNOLOGIES INC.
 XX
 PI Pillutla RC, Brissette R, Spruyt M, Dedova O, Blume A;
 PI Prendergast J, Goldstein N;
 XX
 DR WPI; 2003-457332/43.
 XX

PT Selecting target and target binder pairs for preparing a composition for
 PT treating cancer by mixing in a reaction vessel phase expressing
 PT biological targets and phase expressing target binders.
 XX
 PS Claim 26; SEQ ID NO 476; 172pp; English.

CC The invention relates to a novel method of selecting target and target
 CC binder pairs comprising mixing in a reaction vessel phase expressing
 CC biological targets and phase expressing target binders, each having
 CC distinguishable selection markers and selecting target and target binder
 CC pairs based on the selection markers. The molecules of the invention
 CC demonstrate cytostatic activity whilst the method may be useful for
 CC selecting target and target binder pairs for preparing a composition for
 CC treating cancer. Furthermore, the method may be utilised during gene
 CC therapy procedures. The current sequence is that of the cancer-related
 CC Tie1-binder peptide of the invention.
 XX
 SQ Sequence 25 AA;

Query Match 65.3%; Score 47; DB 7; Length 25;
 Best Local Similarity 66.7%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYWLTIWGC 9
 |||||
 Db 5 CGYWGELWG 13

RESULT 6
 AAO26074
 ID AAO26074 standard; peptide; 17 AA.
 XX
 AC AAO26074;
 XX
 DT 03-APR-2003 (first entry)
 XX
 DE Fc region binding peptide SEQ ID No 54.
 XX

KW Immunoglobulin Fc region; binding; whole blood; plasma; transgenic milk;
 KW antibody response; half-life; stability; circulatory system.
 XX Unidentified.
 XX WO200286070-A2.
 XX
 XX 31-OCT-2002.
 XX
 XX 18-APR-2002; 2002WO-US012492.
 XX
 XX 18-APR-2001; 2001US-0284534P.
 XX
 XX (DYAX-) DYAX CORP.
 XX
 XX Rondon IJ, Wu Q, Ley AC, Stochl M, Ransohoff TC, Potter MD;
 XX WPI; 2003-201220/19.
 XX
 XX New polypeptides, useful as binding molecules for detecting, isolating or
 PT purifying immunoglobulin Fc-region polypeptides present in a solution, or
 PT for regulating or preventing an antibody response.
 XX
 XX Claim 3; Page 76; 152pp; English.
 XX
 XX The invention relates to novel isolated polypeptides comprising a
 CC sequence that binds an immunoglobulin Fc region. The polypeptides are
 CC useful as binding molecules for detecting, isolating or purifying
 CC immunoglobulin Fc-region polypeptides present in a solution, e.g. whole
 CC blood, plasma or transgenic milk. The Fc-region binding polypeptides are
 CC also useful for regulating or preventing an antibody response, or for
 CC increasing the half-life and over all stability of a therapeutic or
 CC diagnostic compound that is administered to or enters the circulatory
 CC system of an individual. This sequence represents an Fc region binding
 CC peptide of the invention
 XX
 XX Sequence 17 AA;
 XX
 XX Query Match 63.9%; Score 46; DB 6; Length 17;
 XX Best Local Similarity 66.7%; Pred. No. 13;
 XX Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 XX
 XX QY 1 CGYWLTIWG 9
 XX ||:|||||
 XX Db 4 CGFWPRIW 12
 XX
 XX RESULT 7
 XX ADJ50741
 XX ID ADJ50741 standard; peptide; 17 AA.
 XX AC ADJ50741;
 XX
 XX 06-MAY-2004 (first entry)
 XX
 XX Human serum albumin binding peptide, Seq ID No 278.
 XX human serum albumin; HSA; serum; blood; tumour; human.
 XX
 XX Homo sapiens.
 XX
 XX WO2003106493-A1.
 XX
 XX 24-DEC-2003.
 XX
 XX 16-JUN-2003; 2003WO-US018896.
 XX
 XX 14-JUN-2002; 2002US-0388642P.
 XX
 XX (DYAX-) DYAX CORP.
 XX
 XX Sato AK, Dawson BM;
 XX
 XX New isolated polynucleotides and polypeptides, useful for treating, e.g.

DR WPI; 2004-082161/08.
 XX
 XX Evaluating sample comprising soluble serum protein by forming complex
 PT comprising serum protein and physically associated compounds using
 PT peptide ligand that specifically binds with proteins, which is separated
 PT and evaluated.
 XX
 XX Disclosure; SEQ ID NO 278; 191pp; English.
 XX
 XX The invention relates to a method of evaluating sample by providing a
 CC soluble serum protein (I), one or more compounds physically associated
 CC with (I), and a (I)-binding agent that comprises a peptide that
 CC specifically binds to (I), allowing the (I)-binding agent to bind to (I)
 CC to form a complex including one or more compounds physically associated
 CC with (I), separating the complex from one or more components of the
 CC sample, and evaluating one or more of the physically associated
 CC compounds. The sample comprises blood or serum, or is obtained from a
 CC biopsy. The sample may also be obtained from a tumour or a region within
 CC 5 mm of a tumour. The method is useful for detecting modulators that
 CC modulate interaction of serum protein-binding compound and serum protein
 CC and for identifying binding ligands for serum protein. The present
 CC sequence represents a serum albumin-binding peptide identified using the
 CC method of the invention.
 XX
 XX Sequence 17 AA;
 XX
 XX Query Match 63.9%; Score 46; DB 8; Length 17;
 XX Best Local Similarity 66.7%; Pred. No. 13;
 XX Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 XX
 XX QY 1 CGYWLTIWG 9
 XX ||:|||||
 XX Db 4 CGFWPRIW 12
 XX
 XX RESULT 8
 XX ADM87650
 XX ID ADM87650 standard; protein; 136 AA.
 XX AC ADM87650;
 XX
 XX 03-JUN-2004 (first entry)
 XX
 XX Human EST derived amino acid sequence SEQ ID NO:743.
 XX
 XX respiratory; cytostatic; antiarthritic; antiinflammatory;
 KW gastrointestinal; antibacterial; immunosuppressive; antidiabetic;
 KW antirheumatic; gene therapy; molecular weight marker; chromosome marker;
 KW chromosome tag; genetic fingerprinting; nutritional supplement; cancer;
 KW inflammatory condition; arthritis; inflammatory bowel disease;
 KW Crohn's disease; sepsis; rheumatoid arthritis; diabetes mellitus type 1;
 KW graft versus host disease; human; expressed sequence tag; EST.
 XX
 XX Homo sapiens.
 XX
 XX WO2004009834-A2.
 XX
 XX 29-JAN-2004.
 XX
 XX 19-JUL-2002; 2002WO-US022858.
 XX
 XX 21-JUL-2001; 2001US-0306971P.
 XX
 XX 28-MAR-2002; 2002US-00112944.
 XX
 XX (NUVE-) NUVELO INC.
 XX
 XX Tang YT, Yang Y, Weng G, Zhang J, Ren F, Xue A, Wang J;
 PI Wehrman T, Ghosh MJ, Wang D, Zhao QA, Wang Z;
 XX
 XX WPI; 2004-143291/14.
 XX
 XX N-PSDB; ADM87432.
 XX
 XX New isolated polynucleotides and polypeptides, useful for treating, e.g.

PT cancer, lung or liver fibrosis, arthritis, inflammatory bowel disease,
 PT Crohn's disease, rheumatoid arthritis, diabetes mellitus type 1 or graft
 PT versus host disease.

PS Example 2; SEQ ID NO 743; 591pp; English.

XX
 CC The present invention describes an isolated polynucleotide (I): (a)
 CC comprising a nucleotide sequence selected from SEQ ID NO:1-244; or (b)
 CC which encodes a polypeptide with biological activity, where the
 CC polynucleotide hybridizes to (I) under stringent hybridisation conditions
 CC or has greater than 99% sequence identity with (I). (I) has respiratory,
 CC cytostatic, antiarthritic, antiinflammatory, gastrointestinal,
 CC antibacterial, immunosuppressive, antidiabetic and antirheumatic
 CC activities, and can be used in gene therapy. (I) can be used for
 CC generating polynucleotides encoding chimeric or fusion proteins and
 CC heterologous protein sequences. The polynucleotides can be used to
 CC express recombinant protein for analysis, characterisation or therapeutic
 CC use; as markers for tissues in which the corresponding protein is
 CC preferentially expressed; as molecular weight markers on gels; as
 CC chromosome markers or tags to identify chromosomes or to map related gene
 CC positions; to compare with endogenous DNA sequences in patients to
 CC identify potential genetic disorders; as probes to hybridise and discover
 CC genes, related DNA sequences; as a source of information to derive PCR
 CC primers for genetic fingerprinting; as a probe to subtract-out known
 CC sequences in the process of discovering other novel polynucleotides; for
 CC selecting and making oligomers for attachment to a gene chip or other
 CC support, including for examination of expression patterns; to raise anti-
 CC protein antibodies using DNA immunisation techniques; and as an antigen
 CC to raise anti-DNA antibodies or elicit another immune response. The
 CC polynucleotides and polypeptides can also be used as nutritional sources
 CC or supplements, e.g. as a protein or amino acid supplement, as a carbon
 CC source, as a nitrogen source or as a source of carbohydrates. The
 CC polynucleotides and polypeptides can also be used treat cancer. The
 CC compositions are useful for promoting better or faster closure of non-
 CC healing wounds, for the generation and regeneration of tissues, for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, and conditions resulting from
 CC systemic cytokine damage. The compositions can also be used to treat
 CC inflammatory conditions (e.g. arthritis, inflammatory bowel disease or
 CC Crohn's disease), sepsis, rheumatoid arthritis, diabetes mellitus type 1
 CC or graft versus host disease. The present sequence represents an
 CC expressed sequence tag (EST) derived amino acid sequence from the present
 CC invention. N.B. The sequences for this patent were obtained from the
 CC USPTO web site from an equivalent US patent US20040048249A1.

XX Sequence 136 AA;

Query Match 63.9%; Score 46; DB 8; Length 136;
 Best Local Similarity 60.0%; Pred. No. 73;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGVWLTWGC 10
 ||| :|||
 DB 93 CGRWDLWGC 102

RESULT 9

AAU90545
 ID AAU90545 standard; peptide; 20 AA.

XX AC AAU90545;

XX DT 18-JUN-2002 (first entry)

XX DE Insulin/insulin-like growth factor receptor-binding peptide #2501.

XX KW Cytostatic; antidiabetic; neuroprotective; cerebroprotective;

XX KW ophthalmological; insulin; receptor; gene therapy; diabetes;

XX KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;

XX KW diabetic retinopathy; neurological diseases; stroke; diabetic neuropathy.

XX OS Synthetic.

XX

PN WO200172771-A2.

XX 04-OCT-2001.

XX PF 29-MAR-2000; 2000WO-US008528.

XX PR 29-MAR-2000; 2000WO-US008528.

XX PA (DGTB-) DGI BIOTECHNOLOGIES LLC.

XX PA (NOVO) NOVO NORDISK AS.

XX PI Beasley J, Blume AJ, Schaeffer L, Pillutla R, Brandt J;

XX PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandecki WS;

XX PI Hansen PH, Ravera M, Heiao K;

XX DR WPI; 2002-025774/03.

XX Modulating insulin activity in mammalian cells, for treating e.g.
 PT diabetes and tumors, comprises using peptides that bind to insulin or
 PT insulin-like growth factor receptors.

XX PS Disclosure; Fig 8-3; 390pp; English.

XX The invention relates to a method of modulating insulin activity in
 CC mammalian cells by administering a peptide that binds the insulin
 CC receptor (IR). A composition containing a peptide, optionally expressed
 CC from gene therapy vectors, that binds to Site 1 of IR and an insulin
 CC agonist are useful for treating diabetes. Also, peptides that are
 CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are
 CC useful for treating insulin-like growth factor (IGF)-sensitive tumours
 CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1
 CC receptor agonists are useful for treating neurological diseases,
 CC including stroke and diabetic neuropathy. The peptides are also useful in
 CC screening for compounds that bind to IR or IGF-1 receptor, potential
 CC therapeutics and research reagents. AAU9034-AAU90357 represent IR and/or
 CC IGF-1 receptor-binding peptides and related amino acid sequences of the
 CC invention

XX Sequence 20 AA;

Query Match 62.5%; Score 45; DB 5; Length 20;
 Best Local Similarity 60.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGVWLTWGC 10
 ||| :|||
 DB 8 CGAWPTWNC 17

RESULT 10

AAU59880
 ID AAU59880 standard; protein; 129 AA.

XX AC AAU59880;

XX DT 19-JAN-2000 (first entry)

XX DE Human normal uterus tissue derived protein 43.

XX KW Human; uterus; cancer; treatment; anticancer; cytostatic; gene therapy;

XX KW EST; expressed sequence tag.

XX OS Homo sapiens.

XX PN DE19817946-A1.

XX PD 21-OCT-1999.

XX PF 17-APR-1998; 98DE-01017946.

XX PR 17-APR-1998; 98DE-01017946.

XX PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

```

XX PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX WPI; 1999-591956/51.
XX DR N-PSDB; AAZ41339.
XX DR
XX PT New nucleic acid sequences expressed in normal uterine tissues, and
XX derived polypeptides, for treatment of uterine cancer and identification
XX of therapeutic agents.
XX PS Claim 23; Page 138; 154pp; German.
XX CC This invention describes novel cDNA sequences (A) highly expressed in
XX normal uterine tissue which can have anticancer and cytostatic activity
XX and can be used for gene therapy. (A) are used (i) for recombinant
XX expression of polypeptides (B) and (ii) to isolate complete genes. (B)
XX are used (i) to identify agents suitable for treatment of uterine cancer;
XX (ii) directly for treating this form of cancer (including expression from
XX gene therapy vectors) and (iii) for generation of specific antibodies.
XX (A) are identified by assembling ESTs (expressed sequence tags) from a
XX particular tissue type before comparison of expression patterns. This
XX allows a significantly longer fragment of the gene to be revealed, so
XX should reduce the number of failures associated with the fact that ESTs
XX from different libraries may represent different parts of the same
XX unknown gene, distorting the estimated frequency of occurrence in a
XX particular tissue. AAY59838-Y59892 represent protein fragments encoded by
XX the human uterine tissue derived cDNA fragments represented in AAZ41325-
XX Z41385
XX SQ Sequence 129 AA;

Query Match 62.5%; Score 45; DB 2; Length 129;
Best Local Similarity 62.5%; Pred. No. 96;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGYWLTIW 8
Db 32 CSHLTIW 39
| : ||| : |
: ||| : |

RESULT 11
ABU49938
ID ABU49938 standard; protein; 828 AA.
XX AC ABU49938;
XX 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #35465.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Yersinia pestis.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITPA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA53808.

```

```

XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 77862; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway;
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 828 AA;

Query Match 62.5%; Score 45; DB 6; Length 828;
Best Local Similarity 58.3%; Pred. No. 4.6e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 1 CGYWLTI--WGC 10
Db 372 CGLWLELLSWGC 383
| | | | : | | |
| | | | : | | |

RESULT 12
ABB04861
ID ABB04861 standard; protein; 1024 AA.
XX AC ABB04861;
XX 13-MAR-2002 (first entry)
XX DE LDL receptor binding protein Na channel brain 3 SEQ ID NO:84.
XX KW Low density lipoprotein receptor binding protein; signal transduction;
XX LDL receptor binding protein; LDL receptor signalling pathway.
XX OS Synthetic.
XX PN WO200184159-A2.
XX PD 08-NOV-2001.
XX PF 24-APR-2001; 2001WO-US013214.
XX PR 01-MAY-2000; 2000US-00562737.
XX

```

PA (TEXA) UNIV TEXAS SYSTEM.
 XX Herz J, Gotthardt M;
 XX WPI; 2002-082855/11.
 XX
 PT Detecting stress that alters interaction of LDL receptor binding
 PT polypeptide with LDL receptor interaction domain, comprises detecting
 PT difference in stress-biased and unbiased interaction of peptide and
 PT domain in a system.
 XX
 XX Disclosure; Page 148-150; 200pp; English.
 XX
 CC The present invention describes a method for detecting a stress that
 CC alters a functional interaction of a low density lipoprotein (LDL)
 CC receptor binding protein (I) with an LDL receptor interaction domain
 CC (II). The method involves introducing a predetermined stress into a
 CC system which provides a stress-biased physical interaction of (I) with
 CC (II), where in the absence of the stress, the system provides an unbiased
 CC interaction of (I) and (II), and detecting the stress-biased interaction
 CC of (I) and (II), where a difference between BI and UI indicates that the
 CC stress alters the interaction of (I) and (II). (I) is selected from
 CC SEMCAP-1, JIP-1, PSD-95, JIP-2, Talin, OMP25, CAPON, PIP4.5 Kinase, Na
 CC channel brain 3, Mint1, ICAP-1 and APC subunit 10. The method is useful
 CC for detecting a stress that alters functional interaction of LDL receptor
 CC binding polypeptide with LDL receptor interaction domain. The method is
 CC useful for detecting and modulating signal transduction through LDL
 CC receptors. ABB04778 to ABB04909 represent LDL receptor binding proteins
 CC which are used in the exemplification of the present invention
 XX
 SQ Sequence 1024 AA;
 Query Match 61.8%; Score 44.5; DB 5; Length 1024;
 Best Local Similarity 54.5%; Pred. No. 6.5e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;
 QY 1 CGYWL-TWGC 10
 ||:|:|:|:
 Db 892 CGHWIETWDC 902
 RESULT 13
 ADP29850
 ID ADP29850 standard; protein; 149 AA.
 XX
 AC ADP29850;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human secreted protein SEQ ID #617.
 XX
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
 KW cancer; inflammatory; immune; human secreted protein.
 XX
 OS Homo sapiens.
 XX
 XX WO2004035732-A2.
 PN
 PD 29-APR-2004.
 XX
 PF 28-AUG-2003; 2003WO-US026780.
 XX
 PR 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406585P.
 PR 29-AUG-2002; 2002US-0406588P.
 PR 29-AUG-2002; 2002US-0406608P.
 PR 29-AUG-2002; 2002US-0406611P.
 PR 29-AUG-2002; 2002US-0406612P.
 PR 29-AUG-2002; 2002US-0406616P.
 PR 29-AUG-2002; 2002US-0406640P.
 PR 29-AUG-2002; 2002US-0406642P.
 PR 29-AUG-2002; 2002US-0406646P.
 PR
 PR (FIVE-) FIVE PRIME THERAPEUTICS INC.
 PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
 PI Halenbeck RP, Huang MM, Kothakota S, Haishan L, Linnemann T;
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
 XX WPI; 2004-348438/32.
 DR
 XX New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.
 XX
 PS Claim 1; SEQ ID NO 1848; 428pp; English.
 XX
 CC The present invention relates to an isolated nucleic acid molecule

CC encoding a polypeptide which is believed to be cytostatic,
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPWEB and is not in the specification.

XX
 SQ Sequence 149 AA;

Query Match 61.1%; Score 44; DB 8; Length 149;
 Best Local Similarity 62.5%; Pred. No. 1.5e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGWLTITW 8
 DB 95 CHFELTW 102

RESULT 14
 AAEE31487
 ID AAE31487 standard; protein; 166 AA.
 XX AC
 XX AAE31487;
 XX
 DT 24-FEB-2003 (first entry)
 XX
 DE Human butyrophilin 2/3 protein.
 XX
 KW Human; B7-H1.2 protein; Butyrophilin 2/3 protein; transplant rejection;
 KW immunological condition; graft-versus-host disease; allergy; asthma;
 KW inflammatory bowel disease; sepsis; Alzheimer's disease; atherosclerosis;
 KW T-cell mediated inflammation; autoimmune disease; multiple sclerosis;
 KW systemic lupus erythematosus; autoimmune demyelination; Grave's disease;
 KW psoriasis; autoimmune diabetes; diabetic neuropathy; HIV infection;
 KW rheumatoid arthritis; human immunodeficiency virus; immunosuppressive;
 KW gene therapy; infection; virucide.
 XX
 OS Homo sapiens.
 XX
 PN WO200279474-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 08-JAN-2002; 2002WO-US0000590.
 XX
 PR 08-JAN-2001; 2001US-0260617P.
 PR 19-JAN-2001; 2001US-0262737P.
 PR 07-JAN-2002; 2002US-0026061P.
 XX
 PA (IMMUNEX) IMMUNEX CORP.
 XX
 XX Baum PR, Dubose RF, Wiley SR;
 XX WPI; 2003-046816/04.
 DR
 XX
 PT New B7-H1.2 or Butyrophilin 2/3 polypeptide of the human B7 polypeptide
 PT family, useful for treating an immunological condition e.g. transplant
 PT rejection.
 PS
 XX
 PS Example 1; Page 64-65; 99pp; English.
 XX
 CC The invention relates to B7-H1.2 or Butyrophilin 2/3 polypeptides of
 CC human B7 polypeptide family and polynucleotides encoding such proteins.
 CC Sequences of the invention are useful for treating an immunological
 CC conditions (e.g., transplant rejection, graft-versus-host disease,
 CC allergy, asthma, inflammatory bowel disease, sepsis), diseases that are
 CC caused or exacerbated by T-cell mediated inflammation (e.g., Alzheimer's
 CC disease or atherosclerosis), autoimmune diseases (e.g., systemic lupus
 CC erythematosus, autoimmune demyelination, Grave's disease, psoriasis,
 CC multiple sclerosis, autoimmune diabetes, diabetic neuropathy, rheumatoid
 CC arthritis), bacterial or viral infections such as human immunodeficiency
 CC virus (HIV) infection, delayed reconstitution of T-cells, defects in T-

CC cell or accessory cell function or congenital immunodeficiencies. They
 CC are also used in gene therapy. The present sequence is human butyrophilin
 CC consensus protein
 XX
 SQ Sequence 166 AA;

Query Match 61.1%; Score 44; DB 6; Length 166;
 Best Local Similarity 62.5%; Pred. No. 1.6e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 YMLTIWGC 10
 DB 127 YNEVWGC 134

RESULT 15
 AAU91293
 ID AAU91293 standard; protein; 415 AA.
 XX AC
 XX AAU91293;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Human NOV8 protein.
 XX
 KW Human; NOVX; gene therapy; cardiomyopathy; atherosclerosis; diabetes;
 KW cell signal processing; metabolic pathway modulation; inflammation;
 KW autoimmune disorder; scleroderma; transplantation; allergy;
 KW systemic lupus erythematosus; haemophilia; Alzheimer's disease;
 KW graft versus host disease; Lesch-Nyhan syndrome; periodontitis;
 KW pancreatitis; musculoskeletal disorder; Parkinson's disease;
 KW Huntington's disease; behavioural disorder; pain; obesity; wound healing;
 KW neurodegenerative disorder; neuropsychiatric disorder; hypertension;
 KW growth disorder; reproductive disorder; lung disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200216600-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 27-AUG-2001; 2001WO-US026518.
 XX
 PR 25-AUG-2000; 2000US-0227800P.
 PR 25-AUG-2000; 2000US-0228205P.
 PR 25-AUG-2000; 2000US-0228324P.
 PR 30-AUG-2000; 2000US-0228997P.
 PR 30-AUG-2000; 2000US-0229185P.
 PR 01-SEP-2000; 2000US-0229780P.
 PR 01-SEP-2000; 2000US-0229848P.
 PR 01-SEP-2000; 2000US-0229850P.
 PR 22-JAN-2001; 2001US-0263337P.
 PR 31-JAN-2001; 2001US-0265518P.
 PR 15-MAR-2001; 2001US-0276451P.
 PR 27-MAR-2001; 2001US-0279196P.
 PR 24-AUG-2001; 2001US-00939398.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Gerlach V, Macdougall JR, Smithson G, Stone DJ, Ellerman K;
 XX Spytak KA, Zerbussen BD, Rastelli L, Verney CAM, Patturajan M;
 XX Tchernev VT, Padigaru M, Taupier RJ;
 XX WPI; 2002-292064/33.
 DR N-PSDB; ABK55577.
 XX
 PT New isolated cytoplasmic, nuclear, membrane bound and secreted
 PT polypeptides, termed NOVX, useful for treating inflammation, autoimmune
 PT disorders, hemophilia, Lesch-Nyhan syndrome, pancreatitis,
 PT musculoskeletal disorders.
 XX
 PS Claim 1; Page 108; 245pp; English.
 XX

CC The invention relates to an isolated cytoplasmic, nuclear, membrane bound
 CC or secreted polypeptide, designated NOVX (actually NOV1, 2a, 2b, 3a, 3b,
 CC 4, 5a, 5b, 5c, 5d, 5f, 5g, 5h, 5i, 6, 7 and 8), a variant of NOVX, a
 CC mature form, or a variant of the mature form of NOVX. Also included are a
 CC polynucleotide encoding NOVX (or its complement), a vector comprising the
 CC polynucleotide, a cell comprising the vector, an anti-NOVX antibody,
 CC determining the presence of NOVX in a sample using the antibody,
 CC which binds to NOVX polynucleotide, identifying an agent which binds to
 CC NOVX (including modulators of NOVX). NOVX, the polynucleotide and the
 CC antibody are useful for diagnosing, treating or preventing a NOVX-
 CC associated disorder selected from cardiomyopathy, atherosclerosis,
 CC diabetes, a disorder related to cell signal processing and metabolic
 CC pathway modulation, inflammation, autoimmune disorders, scleroderma,
 CC transplantation, allergies, systemic lupus erythematosus, haemophilia,
 CC graft versus host disease, Alzheimer's disease, stroke, Leach-Nyhan
 CC syndrome, periodontitis, pancreatitis, musculoskeletal disorders,
 CC Parkinson's disease, Huntington's disease, behavioural disorders, pain,
 CC neurodegenerative and neuropsychiatric disorders, hypertension, wound
 CC healing, obesity, growth and reproductive disorders, lung diseases and
 CC many other diseases and disorders listed in the specification. NOVX, the
 CC polynucleotide and the antibody are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g., diagnostic assays, prognostic assays,
 CC monitoring clinical trials and pharmacogenomic), and in methods of
 CC treatment (e.g., therapeutic and prophylactic). NOVX is useful as
 CC immunogen to produce antibodies immunospecific for NOVX, as vaccines to
 CC screen for potential agonist and antagonist compounds, and as bait
 CC protein in a two-hybrid or three-hybrid assay. The polynucleotide is
 CC useful in gene therapy, to express NOVX, to detect NOVX mRNA or a genetic
 CC lesion in a NOVX gene, and to modulate NOVX activity. The vector is
 CC useful for producing non-human transgenic animals. The antibody is useful
 CC for isolating, and purifying NOVX and to monitor protein levels in tissue
 CC as part of a clinical testing procedure. The present sequence represents
 CC a NOVX protein

XX
 SQ Sequence 415 AA;

Query Match 61.1%; Score 44; DB 5; Length 415;

Best Local Similarity 62.5%; Pred. No. 3.5e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGYWLTIW 8

Db 95 CHEFWLTVW 102

Search completed: January 3, 2005, 15:49:01

Job time : 154 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 15:45:27 ; Search time 46 Seconds
(without alignments)
14.417 Million cell updates/sec

Title: SEQ33

Perfect score: 72

Sequence: 1 cgywltiwcg 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44.5	61.8	1024	4	US-09-562-737-84
2	43	59.7	492	4	US-09-107-532A-6945
3	42.5	59.0	1024	4	US-09-562-737-86
4	42.5	59.0	1765	4	US-09-354-147C-2
5	42.5	59.0	1765	4	US-09-354-147C-3
6	42	58.3	151	4	US-09-325-932A-187
7	42	58.3	273	4	US-09-270-767-46926
8	42	58.3	1498	4	US-09-792-616-9
9	42	58.3	1503	4	US-09-792-616-3
10	41	56.9	222	4	US-09-071-035-408
11	41	56.9	229	4	US-09-134-000C-3630
12	41	56.9	264	4	US-09-540-236-2978
13	41	56.9	266	4	US-09-071-035-406
14	40.5	56.2	1024	4	US-09-562-737-81
15	40.5	56.2	1024	4	US-09-562-737-87
16	40.5	56.2	1835	3	US-08-836-325-15
17	40.5	56.2	1835	4	US-09-457-571-15
18	40.5	56.2	1836	4	US-10-162-012-24
19	40.5	56.2	1969	3	US-08-836-325-16
20	40.5	56.2	1969	4	US-09-457-571-16
21	40.5	56.2	1976	3	US-09-024-020B-9
22	40.5	56.2	1976	3	US-09-425-043-9
23	40.5	56.2	1977	4	US-09-976-594-757
24	40.5	56.2	1977	4	US-09-919-039-367
25	40.5	56.2	1978	3	US-09-024-020B-3
26	40.5	56.2	1978	3	US-09-425-043-3
27	40.5	56.2	1984	3	US-08-836-325-10

28	40.5	56.2	1984	4	US-09-457-571-10	Sequence 10, Appl
29	40.5	56.2	1988	3	US-09-024-020B-4	Sequence 4, Appli
30	40.5	56.2	1988	3	US-09-425-043-4	Sequence 4, Appli
31	40.5	56.2	1989	3	US-08-836-325-11	Sequence 11, Appl
32	40.5	56.2	1989	3	US-08-836-325-12	Sequence 12, Appl
33	40.5	56.2	1989	4	US-09-457-571-11	Sequence 11, Appl
34	40.5	56.2	1989	4	US-09-457-571-12	Sequence 12, Appl
35	40.5	56.2	2005	3	US-08-836-325-7	Sequence 7, Appli
36	40.5	56.2	2005	4	US-09-457-571-7	Sequence 7, Appli
37	40.5	56.2	2016	3	US-09-634-920-4	Sequence 4, Appli
38	40.5	56.2	2016	4	US-09-514-907A-2	Sequence 2, Appli
39	40.5	56.2	2016	4	US-09-896-994-2	Sequence 2, Appli
40	40.5	56.2	2016	4	US-09-840-125-4	Sequence 4, Appli
41	40	55.6	362	1	US-08-415-751-6	Sequence 6, Appli
42	40	55.6	668	4	US-09-248-796A-19350	Sequence 19350, A
43	39.5	54.9	73	4	US-09-513-999C-4541	Sequence 4541, Ap
44	39	54.2	430	4	US-09-443-041A-18	Sequence 18, Appl
45	39	54.2	450	4	US-09-443-041A-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-09-562-737-84
; Sequence 84, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 84
; TYPE: PRT
; LENGTH: 1024
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-84

Query Match 61.8%; Score 44.5; DB 4; Length 1024;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10

Db 892 CGHWIETWDC 902

RESULT 2

US-09-107-532A-6945
; Sequence 6945, Application US/09107532A
; Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660

COMPUTER: PC

```
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6945:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 492 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...492
; SEQUENCE DESCRIPTION: SEQ ID NO: 6945:
US-09-107-532A-6945

Query Match 59.7%; Score 43; DB 4; Length 492;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GYWLTIW 8
Db 112 GYWLTCW 118

RESULT 3
US-09-562-737-86
; Sequence 86, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-86

Query Match 59.0%; Score 42.5; DB 4; Length 1024;
Best Local Similarity 63.6%; Pred. No. 3.1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Qy 1 GGYWL-TIWGC 10
Db 892 CGEWLETWDC 902

RESULT 4
US-09-354-147C-2
; Sequence 2, Application US/09354147C
; Patent No. 6573067
; GENERAL INFORMATION:
; APPLICANT: Dib-Hajj, Sulayman
; TITLE OF INVENTION: Modulation of Sodium Channels in Dorsal Root Ganglia
; FILE REFERENCE: 44574-5004-01-US
; CURRENT APPLICATION NUMBER: US/09/354,147C
; CURRENT FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US 60/072,990
; PRIOR FILING DATE: 1998-01-29
; PRIOR APPLICATION NUMBER: US 60/109,402
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: PCT/US99/02008
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1765
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (652)...(1334)
; OTHER INFORMATION: Xaa at position 652 is Leu; Xaa at position 1334 is Asn
; OTHER INFORMATION: or Lys. Xaa's result from n's in SEQ ID NO: 1.
US-09-354-147C-2

Query Match 59.0%; Score 42.5; DB 4; Length 1765;
Best Local Similarity 54.5%; Pred. No. 5.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 CGYWL-TIWGC 10
Db 754 CGEWIENMGC 764

RESULT 5
US-09-354-147C-3
; Sequence 3, Application US/09354147C
; Patent No. 6573067
; GENERAL INFORMATION:
; APPLICANT: Dib-Hajj, Sulayman
; TITLE OF INVENTION: Modulation of Sodium Channels in Dorsal Root Ganglia
; FILE REFERENCE: 44574-5004-01-US
; CURRENT APPLICATION NUMBER: US/09/354,147C
; CURRENT FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US 60/072,990
; PRIOR FILING DATE: 1998-01-29
; PRIOR APPLICATION NUMBER: US 60/109,402
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: PCT/US99/02008
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1765
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: putative amino acid seq. of rat Nan
US-09-354-147C-3

Query Match 59.0%; Score 42.5; DB 4; Length 1765;
Best Local Similarity 54.5%; Pred. No. 5.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 CGYWL-TIWGC 10
Db 754 CGEWIENMGC 764
```



```
RESULT 6
US-09-325-932A-187
; Sequence 187, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of forestry plant develop
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 187
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-325-932A-187

Query Match      58.3%; Score 42; DB 4; Length 161;
Best Local Similarity 66.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GYWLTIWGC 10
Db      100 GYWLNSGC 108
      |||||: ||
      |||||: ||

RESULT 7
US-09-270-767-46926
; Sequence 46926, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46926
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46926

Query Match      58.3%; Score 42; DB 4; Length 273;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 CGYWLTIW 8
Db      266 CGYWTATF 273
      |||||: ||
      |||||: ||

RESULT 8
US-09-792-616-9
; Sequence 9, Application US/09792616
; Patent No. 6780587
; GENERAL INFORMATION:
; APPLICANT: PXE International, Inc.
; APPLICANT: University of Hawaii
; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (MRP6) causing
; TITLE OF INVENTION: Pseudoxanthoma Elasticum
; FILE REFERENCE: PXE-001
; CURRENT APPLICATION NUMBER: US/09/792,616
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1498
```

```
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-792-616-9

Query Match      58.3%; Score 42; DB 4; Length 1498;
Best Local Similarity 71.4%; Pred. No. 5.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 GYWLTIW 8
Db      960 GYWLSLW 966
      |||||: ||
      |||||: ||

RESULT 9
US-09-792-616-3
; Sequence 3, Application US/09792616
; Patent No. 6780587
; GENERAL INFORMATION:
; APPLICANT: PXE International, Inc.
; APPLICANT: University of Hawaii
; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (MRP6) causing
; TITLE OF INVENTION: Pseudoxanthoma Elasticum
; FILE REFERENCE: PXE-001
; CURRENT APPLICATION NUMBER: US/09/792,616
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1503
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-616-3

Query Match      58.3%; Score 42; DB 4; Length 1503;
Best Local Similarity 71.4%; Pred. No. 5.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 GYWLTIW 8
Db      965 GYWLSLW 971
      |||||: ||
      |||||: ||

RESULT 10
US-09-071-035-408
; Sequence 408, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 408:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-408

Query Match 56.9%; Score 41; DB 4; Length 222;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GYWLTIW 9
Db 179 GTWITLWG 186

RESULT 11
US-09-134-000C-3630
; Sequence 3630, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3630
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3630

Query Match 56.9%; Score 41; DB 4; Length 229;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GYWLTIW 9
Db 203 GTWITLWG 210

RESULT 12
US-09-540-236-2978
; Sequence 2978, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2978
; LENGTH: 264
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2978

Query Match 56.9%; Score 41; DB 4; Length 264;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 408:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-406

Query Match 56.9%; Score 41; DB 4; Length 266;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GYWLTIW 9
Db 203 GTWITLWG 210

RESULT 13
US-09-071-035-406
; Sequence 406, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB3699P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 406:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 266 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-406

Query Match 56.9%; Score 41; DB 4; Length 266;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GYWLTIW 9
Db 203 GTWITLWG 210

RESULT 14
US-09-562-737-81
; Sequence 81, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 1024
; TYPE: PRT
```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-81

Query Match 56.2%; Score 40.5; DB 4; Length 1024;
Best Local Similarity 54.5%; Pred. No. 5.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10
|||:|:|
DB 892 CGEWIGTMWDC 902

RESULT 15
US-09-562-737-87
; Sequence 87, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-87

Query Match 56.2%; Score 40.5; DB 4; Length 1024;
Best Local Similarity 54.5%; Pred. No. 5.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10
|||:|:|
DB 892 CGEWIGTMWDC 902

Search completed: January 3, 2005, 15:53:49
Job time : 47 secs

THIS PAGE BLANK (user's)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 15:52:22 ; Search time 142 seconds
(without alignments)
25.333 Million cell updates/sec

Title: SEQ33

Perfect score: 72

Sequence: 1 cgywlwtgwc 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	100.0	10	13	US-10-046-922-35
2	54	75.0	10	13	US-10-046-922-34
3	50	69.4	10	13	US-10-046-922-73
4	47.5	66.0	304	14	US-10-017-161-2288
5	47.5	66.0	304	14	US-10-292-798-1934
6	47	65.3	25	14	US-10-280-066-476
7	47	65.3	63	16	US-10-437-963-125253
8	46	63.9	17	14	US-10-125-869A-54
9	46	63.9	17	15	US-10-462-262-278
10	46	63.9	136	15	US-10-112-944-743
11	45	62.5	838	15	US-10-282-122A-77862
12	44.5	61.8	1024	14	US-10-211-962-84
13	44	61.1	103	16	US-10-437-963-181150
					Sequence 35, Appl
					Sequence 34, Appl
					Sequence 73, Appl
					Sequence 2288, Ap
					Sequence 1934, Ap
					Sequence 476, App
					Sequence 125253,
					Sequence 54, Appl
					Sequence 278, App
					Sequence 743, App
					Sequence 77862, A
					Sequence 84, Appl
					Sequence 181150,

14	44	61.1	161	16	US-10-767-701-54778	Sequence 54778, A
15	43	59.7	13	14	US-10-125-869A-73	Sequence 73, Appl
16	43	59.7	13	15	US-10-462-262-297	Sequence 297, App
17	43	59.7	291	14	US-10-369-493-19313	Sequence 19313, A
18	43	59.7	454	14	US-10-032-585-7816	Sequence 7816, Ap
19	43	59.7	459	16	US-10-437-963-160191	Sequence 160191,
20	43	59.7	474	15	US-10-282-122A-57928	Sequence 57928, A
21	43	59.7	671	16	US-10-437-963-190740	Sequence 190740,
22	42.5	59.0	152	11	US-09-833-245-1533	Sequence 1533, Ap
23	42.5	59.0	152	15	US-10-264-237-2457	Sequence 2457, Ap
24	42.5	59.0	159	11	US-09-833-245-1534	Sequence 1534, Ap
25	42.5	59.0	161	15	US-10-276-774-2337	Sequence 2376, Ap
26	42.5	59.0	161	15	US-10-296-115-1337	Sequence 1337, Ap
27	42.5	59.0	176	14	US-10-161-927-10	Sequence 10, Appl
28	42.5	59.0	1024	14	US-10-211-962-86	Sequence 86, Appl
29	42.5	59.0	1765	14	US-10-388-470-2	Sequence 2, Appl
30	42.5	59.0	1765	14	US-10-388-470-3	Sequence 3, Appl
31	42	58.3	90	11	US-09-864-408A-8146	Sequence 8146, Ap
32	42	58.3	161	14	US-10-219-220-187	Sequence 187, App
33	42	58.3	1498	10	US-09-792-616-9	Sequence 9, Appl
34	42	58.3	1498	16	US-10-764-328-9	Sequence 9, Appl
35	42	58.3	1503	10	US-09-792-616-3	Sequence 3, Appl
36	42	58.3	1503	16	US-10-764-328-3	Sequence 3, Appl
37	41.5	57.6	122	15	US-10-108-260A-2641	Sequence 2641, Ap
38	41.5	57.6	620	14	US-10-369-493-119	Sequence 119, App
39	41	56.9	14	14	US-10-125-869A-114	Sequence 114, App
40	41	56.9	14	15	US-10-462-262-338	Sequence 338, App
41	41	56.9	34	9	US-09-864-761-43458	Sequence 43458, A
42	41	56.9	57	15	US-10-424-599-179308	Sequence 179308,
43	41	56.9	78	17	US-10-425-115-204642	Sequence 204642,
44	41	56.9	82	15	US-10-424-599-240891	Sequence 240891,
45	41	56.9	104	17	US-10-425-115-332065	Sequence 332065,

ALIGNMENTS

RESULT 1

US-10-046-922-35
; Sequence 35, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 35
; LENGTH: 10
; TYPE: PRT
; ORGANISM: isolated peptide
US-10-046-922-35

Query Match 100.0%; Score 72; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGVWLTIWGC 10

Db 1 CGVWLTIWGC 10

RESULT 2

US-10-046-922-34
; Sequence 34, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki

```
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 10
; TYPE: PRT
; ORGANISM: isolated peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(1)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (10)..(10)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-34

Query Match 75.0%; Score 54; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GYWLTIWG 9
Db 2 GYWLTIWG 9

RESULT 3
US-10-046-922-73
; Sequence 73, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73
; LENGTH: 10
; TYPE: PRT
; ORGANISM: peptide library
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)..(7)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (9)..(9)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-73

Query Match 69.4%; Score 50; DB 13; Length 10;
Best Local Similarity 60.0%; Pred. No. 2.5;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGYWLTIWGC 10
Db 1 CGYWXXXWC 10

RESULT 4
US-10-017-161-2288
; Sequence 2288, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
```

```
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2288
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (73)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (79)..(83)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (85)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (89)..(91)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (96)..(97)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (101)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (107)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (118)..(119)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (121)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (123)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (162)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (178)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (194)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (210)..(211)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (219)
; OTHER INFORMATION: Variable amino acid
```

```
/
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (256)..(289)
; OTHER INFORMATION: Variable amino acid
US-10-017-161-2288

Query Match          66.0%; Score 47.5; DB 14; Length 304;
Best Local Similarity 60.0%; Pred. No. 88;
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY      1 CGWLTWGC 10
DB      108 CGFW-AWGC 116

RESULT 5
US-10-292-798-1934
; Sequence 1934, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1934
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (73)..(73)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (80)..(83)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (89)..(91)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (96)..(97)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (101)..(101)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (107)..(107)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (118)..(119)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (121)..(121)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (123)..(123)
```

```
/
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (178)..(178)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (210)..(211)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (219)..(219)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (256)..(289)
; OTHER INFORMATION: Variable amino acid
US-10-292-798-1934

Query Match          66.0%; Score 47.5; DB 14; Length 304;
Best Local Similarity 60.0%; Pred. No. 88;
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY      1 CGWLTWGC 10
DB      108 CGFW-AWGC 116

RESULT 6
US-10-280-066-476
; Sequence 476, Application US/10280066
; Publication No. US20030180718A1
; GENERAL INFORMATION:
; APPLICANT: Pillutla, Renuka C.
; APPLICANT: Brissette, Renee
; APPLICANT: Spruyt, Michael
; APPLICANT: Dedova, Olga
; APPLICANT: Blume, Arthur J.
; APPLICANT: Prendergast, John
; APPLICANT: Goldstein, Neil I.
; TITLE OF INVENTION: TARGET SPECIFIC SCREENING AND ITS USE FOR IDENTIFYING TARGET BIND
; FILE REFERENCE: 2598-4009US1
; CURRENT APPLICATION NUMBER: US/10/280,066
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 60/345,471
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 537
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 476
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Eschericia coli
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: Tiel-20C-3-D116
US-10-280-066-476

Query Match          65.3%; Score 47; DB 14; Length 25;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 CGWLTWGC 9
DB      5 CGWGTGELWG 13

RESULT 7
US-10-437-963-125253
; Sequence 125253, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
```

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 125253
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_27915C.1.pap
US-10-437-963-125253

Query Match 65.3%; Score 47; DB 16; Length 63;
Best Local Similarity 60.0%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGYWLTIWGC 10
Db 37 CGHYLKAWC 46

RESULT 8
US-10-125-869A-54
; Sequence 54, Application US/10125869A
; Publication No. US20030199671A1
; GENERAL INFORMATION:
; APPLICANT: Rondon, Isaac Jesus
; APPLICANT: Wu, Qi-Long
; APPLICANT: Ley, Arthur C.
; APPLICANT: Stochl, Mark
; APPLICANT: Ransohoff, Thomas C.
; APPLICANT: Potter, M. Daniel (deceased)
; TITLE OF INVENTION: BINDING MOLECULES FOR FC-REGION
; FILE REFERENCE: 3421.1006-001
; CURRENT APPLICATION NUMBER: US/10/125,869A
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/284,534
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fc region binding polypeptide
US-10-125-869A-54

Query Match 63.9%; Score 46; DB 14; Length 17;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGYWLTIW 9
Db 4 CGFWPRWG 12

RESULT 9
US-10-462-262-278
; Sequence 278, Application US/10462262
; Publication No. US20040009534A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Aaron K.
; APPLICANT: Dawson, Bruce M.
```

```
; TITLE OF INVENTION: PROTEIN ANALYSIS
; FILE REFERENCE: 10280-052001
; CURRENT APPLICATION NUMBER: US/10/462,262
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/388,642
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 278
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: immunoglobulin binding polypeptide
US-10-462-262-278

Query Match 63.9%; Score 46; DB 15; Length 17;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGYWLTIW 9
Db 4 CGFWPRWG 12

RESULT 10
US-10-112-944-743
; Sequence 743, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Weng, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; TITLE OF INVENTION: Secreted Polypeptides
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 743
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(136)
; OTHER INFORMATION: Xaa = any amino acid or symbol as shown in the table 8 as set for
```



```
; OTHER INFORMATION: in Example 2
US-10-112-944-743

Query Match      63.9%; Score 46; DB 15; Length 136;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 CGYWLTIWGC 10
Db      93 CGRWDWLWGC 102

RESULT 11
US-10-282-122A-77862
; Sequence 77862, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77862
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Yersinia pestis
US-10-282-122A-77862

Query Match      62.5%; Score 45; DB 15; Length 828;
Best Local Similarity 58.3%; Pred. No. 4.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY      1 CGYWLTI--WGC 10
Db      372 CGLWLELLSNGC 383

RESULT 12
US-10-211-962-84

; Sequence 84, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562,737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-10-211-962-84

Query Match      61.8%; Score 44.5; DB 14; Length 1024;
Best Local Similarity 54.5%; Pred. No. 6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY      1 CGYWL-TIWGC 10
Db      892 CGHWIETMWDG 902

RESULT 13
US-10-437-963-181150
; Sequence 181150, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(S3221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 181150
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_78451C.1.pep
US-10-437-963-181150

Query Match      61.1%; Score 44; DB 16; Length 103;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGYWLTIWGC 10
Db      19 WWVSVWGC 26

RESULT 14
US-10-767-701-54778
; Sequence 54778, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
```

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 54778
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 14593586.pep
US-10-767-701-54778

Query Match 61.1%; Score 44; DB 16; Length 161;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGYWLTIW 8
||| |
Db 150 CGYWLKFW 157

RESULT 15
US-10-125-869A-73
; Sequence 73, Application US/10125869A
; Publication No. US20030199671A1
; GENERAL INFORMATION:
; APPLICANT: Rondon, Isaac Jesus
; APPLICANT: Wu, Qi-Long
; APPLICANT: Lev, Arthur C.
; APPLICANT: Stochl, Mark
; APPLICANT: Ransohoff, Thomas C.
; APPLICANT: Potter, M. Daniel (deceased)
; TITLE OF INVENTION: BINDING MOLECULES FOR FC-REGION
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 3421.1006-001
; CURRENT APPLICATION NUMBER: US/10/125,869A
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/284,534
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fc region binding polypeptide
US-10-125-869A-73

Query Match 59.7%; Score 43; DB 14; Length 13;
Best Local Similarity 62.5%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GYWLTIWG 9
||| :||
Db 1 GYWCNFWG 8

Search completed: January 3, 2005, 16:04:48
Job time : 153 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 15:41:46 ; Search time 38 Seconds
(without alignments)
25.320 Million cell updates/sec

Title: SEQ33

Perfect score: 72

Sequence: 1 cgywltiwcg 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	63.9	218	2 S76385	hypothetical prote
2	45	62.5	828	2 AD0412	ATP-dependent heli
3	43.5	60.4	1993	2 T30902	sodium channel sca
4	42.5	59.0	1765	2 T42388	sodium channel alp
5	42	58.3	72	2 S30980	gene 35 protein -
6	42	58.3	441	2 C95307	probable transport
7	42	58.3	1502	2 T42216	multidrug resistan
8	41.5	57.6	620	2 F70439	oxaloacetate decar
9	41	56.9	307	2 T48166	hypothetical prote
10	41	56.9	391	2 FC4117	replication protei
11	41	56.9	419	2 E90446	permease [imported
12	41	56.9	739	2 T29407	hypothetical prote
13	40.5	56.2	200	2 I48108	sodium channel alp
14	40.5	56.2	1681	2 A55138	sodium channel mna
15	40.5	56.2	1682	2 A45380	sodium channel pro
16	40.5	56.2	1820	1 CHEE	sodium channel pro
17	40.5	56.2	1835	2 I54323	sodium channel alp
18	40.5	56.2	1836	2 I64893	sodium channel alp
19	40.5	56.2	1836	2 J50648	sodium channel alp
20	40.5	56.2	1836	2 I51964	sodium channel alp
21	40.5	56.2	1840	1 CHRTM1	sodium channel pro
22	40.5	56.2	1951	2 S00320	sodium channel pro
23	40.5	56.2	1976	2 I56555	sodium channel pro
24	40.5	56.2	1977	2 I54771	sodium channel alp
25	40.5	56.2	1983	2 A60054	sodium channel pro
26	40.5	56.2	2005	2 A46269	sodium channel alp
27	40.5	56.2	2005	2 B25019	sodium channel pro
28	40.5	56.2	2009	2 A25019	sodium channel pro
29	40.5	56.2	2016	2 A38195	sodium channel pro

30	40.5	56.2	2019	2 A33996	sodium channel pro
31	40.5	56.2	2049	2 T43161	sodium channel pro
32	40	55.6	142	2 C34903	Ig heavy chain pre
33	40	55.6	339	2 F97190	phenylalanyl-tRNA
34	40	55.6	359	2 F95406	probable ABC trans
35	40	55.6	425	2 B71038	probable Na+/H+-ex
36	40	55.6	508	2 C95282	probable ABC trans
37	40	55.6	2344	2 S64740	genome polyprotein
38	39.5	54.9	298	2 AH0289	probable aldo/keto
39	39.5	54.9	345	1 JH0185	D-amino-acid oxida
40	39.5	54.9	347	1 OXPGDA	D-amino-acid oxida
41	39.5	54.9	347	1 S01340	D-amino-acid oxida
42	39.5	54.9	347	1 JX0132	hypothetical prote
43	39	54.2	376	2 AF1978	neurexin III beta
44	39	54.2	392	2 A53580	neurexin III beta
45	39	54.2	426	2 B53580	neurexin III beta

ALIGNMENTS

RESULT 1

S76385

hypothetical protein - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C:Accession: S76385

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

s.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S76385

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-218 <XAN>

A:Cross-references: UNIPROT:Q55705; EMBL:D64000; GB:AB001339; NID:gl001484; PIDN:BAAL023

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Superfamily: probable alkaline phosphatase yngC

Query Match 63.9%; Score 46; DB 2; Length 218;
Best Local Similarity 66.7%; Pred. No. 8.1;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYWLTIWG 9

DB 73 CGYWVGWRWG 81

RESULT 2

AD0412

ATP-dependent helicase [imported] - *Yersinia pestis* (strain CO92)

C:Species: *Yersinia pestis*

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.;

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, J.

Nature 413, 523-527, 2001

A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AD0412

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-828 <KUR>

A:Cross-references: UNIPROT:Q8ZBL4; GB:AL590842; PIDN:CAC92624.1; PID:gl5981320; GSPDB:G

C:Genetics:

A:Gene: hrpB

C:Superfamily: ATP-dependent RNA helicase, HrpB type

Query Match 62.5%; Score 45; DB 2; Length 828;

Best Local Similarity 58.3%; Pred. No. 36;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 CGYWLTI--WGC 10
||| : |||
Db 372 CGLWLELLSWG 383

RESULT 3
T30902
sodium channel SCAP1 alpha chain - California sea hare
C;Species: Aplysia californica (California sea hare)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30902
R;Dyer, J.R.; Johnston, W.L.; Castellucci, V.F.; Dunn, R.J.
DNA Cell Biol. 16, 347-356, 1997
A;Title: Cloning and tissue distribution of the Aplysia Na+ channel alpha-subunit cDNA.
A;Reference number: Z20929; MUID:97238630; PMID:9115644
A;Accession: T30902
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1993 <BYE>
A;Cross-references: UNIPROT:P90670; EMBL:U66915; PID:g1842248; PID:g1842249; PIDN:AAC474
C;Superfamily: sodium channel protein

Query Match 60.4%; Score 43.5; DB 2; Length 1993;
Best Local Similarity 54.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10
||| : |||
Db 942 CGEWIESMWGC 952

RESULT 4
T42388
sodium channel alpha chain - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T42388
R;Dib-Hajj, S.D.; Tyrell, L.; Black, J.A.; Waxman, S.G.
Proc. Natl. Acad. Sci. U.S.A. 95, 8963-8968, 1998
A;Title: NaN, a novel voltage-gated Na channel, is expressed preferentially in peripheral
A;Reference number: Z22149; MUID:98338024; PMID:9671787
A;Accession: T42388
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1765 <DIB>
A;Cross-references: UNIPROT:O88457; EMBL:AF059030; NID:g3372614; PID:g3372615; PIDN:AAC4
A;Experimental source: strain Sprague-Dawley; dorsal root ganglia
A;Note: preferentially expressed in sensory neurons within dorsal root ganglia and trigem
C;Superfamily: sodium channel protein

Query Match 59.0%; Score 42.5; DB 2; Length 1765;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10
||| : |||
Db 754 CGEWIENMWGC 764

RESULT 5
S30980
gene 35 protein - Mycobacterium phage L5
C;Species: Mycobacterium phage L5
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: S30980
R;Donnelly-Wu, M.K.; Jacobs Jr., W.R.; Hatfull, G.F.
Mol. Microbiol. 7, 407-417, 1993
A;Title: Superinfection immunity of mycobacteriophage L5: applications for genetic trans
A;Reference number: S30949; MUID:93211283; PMID:8459767
A;Accession: S30980

A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-72 <DON>
A;Cross-references: UNIPROT:Q05245; EMBL:Z18946; NID:g15859; PIDN:CAA79411.1; PID:g15891
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1992
C;Genetics:
A;Gene: 35

Query Match 58.3%; Score 42; DB 2; Length 72;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYWLTIW 8
||| : |||
Db 50 CGMWLPVW 57

RESULT 6
C95307
probable transport protein Sma0684 [imported] - Sinorhizobium meliloti (strain 1021) mag
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: C95307
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowse
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.;
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: C95307
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-441 <KUR>
A;Cross-references: UNIPROT:Q92ZT6; GB:AE006469; PIDN:AAK65021.1; PID:g14523451; GSPDB:G
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: Sma0684
A;Genome: plasmid
C;Superfamily: L-lysine transport protein

Query Match 58.3%; Score 42; DB 2; Length 441;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GYWLTIW 8
||| : |||
Db 92 GYWISW 98

RESULT 7
T42216
multidrug resistance-associated protein homolog MLP-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T42216
R;Hirohashi, T.; Suzuki, H.; Ito, K.; Ogawa, K.; Kume, K.; Shimizu, T.; Sugiyama, Y.
Mol. Pharmacol. 53, 1068-1075, 1998
A;Title: Hepatic expression of multidrug resistance-associated protein-like proteins main
A;Reference number: Z22081; MUID:98279126; PMID:9614210
A;Accession: T42216
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1502 <HIR>
A;Cross-references: UNIPROT:O88269; EMBL:AB010466; NID:g3242457; PIDN:BAA28954.1; PID:g3242457;
A;Experimental source: strain Sprague-Dawley; liver

C;Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 58.3%; Score 42; DB 2; Length 1502;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GYWLTIW 8
|||:|:
Db 964 GYWLISLW 970

RESULT 8
F70439
oxaloacetate decarboxylase alpha chain - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: F70439
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: F70439
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-620 <AQF>
A;Cross-references: UNIPROT:O67544; GB:AE000747; NID:g2983944; PIDN:AA07497.1; PID:g298
A;Experimental source: strain VF5
C;Genetics:
A;Gene: oadA

C;Superfamily: Klebsiella pneumoniae oxaloacetate decarboxylase alpha chain; lipoyl/biot
F;540-613/Domain: lipoyl/biotin-binding homology <LPB>
F;579/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 57.6%; Score 41.5; DB 2; Length 620;
Best Local Similarity 60.0%; Pred. No. 93;
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 CGYWLTIW 9
|||:|:
Db 38 CGFWSLEWVG 47

RESULT 9
T48166
hypothetical protein T1008.150 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48166
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Lemcke, K.; Mayer, K.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24486
A;Accession: T48166
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-307 <BEV>
A;Cross-references: UNIPROT:Q9M030; EMBL:AL161746
A;Experimental source: cultivar Columbia; EAC clone T1008
C;Genetics:
A;Map position: 5
A;Introns: 31/3; 66/2; 114/3; 149/2; 232/2; 284/1
A;Note: T1008.150

Query Match 56.9%; Score 41; DB 2; Length 307;
Best Local Similarity 50.0%; Pred. No. 60;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYWLTIWGC 10
|:|:|:
Db 186 CSFVSIWGC 195

RESULT 10

PC4117
replication protein homolog - Pyrococcus sp. (fragment)
N;Alternate names: hypothetical 391 protein
C;Species: Pyrococcus sp.
C;Date: 07-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 17-May-1996
C;Accession: PC4117
R;Rashid, N.; Morikawa, M.; Imanaka, T.
Gene 166, 139-143, 1995
A;Title: An abnormally acidic TATA-binding protein from a hyperthermophilic archaeon.
A;Reference number: JC4514; MUID:96105215; PMID:8529878
A;Accession: PC4117
A;Molecule type: DNA
A;Residues: 1-391 <RAS>
A;Cross-references: DDBJ:D50018

Query Match 56.9%; Score 41; DB 2; Length 391;
Best Local Similarity 85.7%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 YWLTIW 9
|||:|:
Db 155 YWLTEWVG 161

RESULT 11
E90446
permease [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: E90446
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: E90446
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-419 <KUR>
A;Cross-references: UNIPROT:Q97VB7; GB:AE006641; NID:gl3916037; PIDN:AAK42828.1; GSPDB:G
C;Genetics:
A;Gene: SSO2718

Query Match 56.9%; Score 41; DB 2; Length 419;
Best Local Similarity 62.5%; Pred. No. 78;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYWLTIW 9
|||:|:
Db 365 GFWEITWVG 372

RESULT 12
T29407
hypothetical protein C16C8.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T29407
R;Waterston, R.; Le, T.T.; Gattung, S.
submitted to the EMBL Data Library, November 1996
A;Description: The sequence of C. elegans cosmid C16C8.
A;Reference number: Z20617
A;Accession: T29407
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-739 <WAT>
A;Cross-references: EMBL:U80452; PIDN:AA37863.1; GSPDB:GN00020; CBSP:C16C8.2
A;Experimental source: strain Bristol N2; clone C16C8
C;Genetics:
A;Gene: CBSP:C16C8.2
A;Map position: 2
A;Introns: 68/2; 179/3; 253/1; 275/3; 327/2; 365/3; 397/1; 428/2; 463/3; 629/1; 668/2; 71

C;Superfamily: myeloperoxidase; myeloperoxidase homology

Query Match 56.9%; Score 41; DB 2; Length 719;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGYWLTIWGC 10
|||:|||||
Db 118 CGFWATIREC 127

RESULT 13

sodium channel alpha subunit - long-tailed hamster (fragment)
C;Species: Crictetus longicaudatus (long-tailed hamster)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I48108
R;Lalick, P.H.; Krafte, D.S.; Ciccarelli, R.B.
Am J. Physiol. 264, 803-809, 1993
A;Title: Characterization of endogenous Sodium channel gene expressed in chinese hamster
A;Reference number: I48107
A;Accession: I48108
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-200 <RES>
A;Cross-references: UNIPROT:Q60464; GB:M87541; NID:g191069; PIDN:AAA36979.1; PID:g553840
C;Genetics:
A;Gene: Chol
C;Superfamily: sodium channel protein
C;Keywords: duplication

Query Match 56.2%; Score 40.5; DB 2; Length 200;
Best Local Similarity 54.5%; Pred. No. 49;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10
|||:|||||
Db 26 CGEWITWDC 36

RESULT 14

A55138
sodium channel mna2.3, voltage-gated - mouse
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: A55138
R;Felipe, A.; Knittle, T.J.; Doyle, K.L.; Tamkun, M.M.
J. Biol. Chem. 269, 30125-30131, 1994
A;Title: Primary structure and differential expression during development and pregnancy
A;Reference number: A55138; MUID:95074002; PMID:7982916
A;Accession: A55138
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1681 <PEL>
A;Cross-references: UNIPROT:Q62467; GB:L36179; NID:g609544; PIDN:AAA66192.1; PID:g806397
C;Superfamily: sodium channel protein

Query Match 56.2%; Score 40.5; DB 2; Length 1681;
Best Local Similarity 54.5%; Pred. No. 3.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10
|||:|||||
Db 686 CGEWITWDC 696

RESULT 15

A45380
sodium channel protein alpha chain hNav2.1 - human
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: A45380
R;George Jr., A.L.; Knittle, T.J.; Tamkun, M.M.

Proc. Natl. Acad. Sci. U.S.A. 89, 4893-4897, 1992
A;Title: Molecular cloning of an atypical voltage-gated sodium channel expressed in human
A;Reference number: A45380; MUID:92279233; PMID:1317577
A;Accession: A45380
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1682 <GEO>
A;Cross-references: UNIPROT:Q01118; GB:M91556; NID:g189046; PIDN:AAA59899.1; PID:g189047
A;Experimental source: heart
A;Note: sequence extracted from NCBI backbone (NCBIP:104344)
C;Superfamily: sodium channel protein
C;Keywords: glycoprotein; membrane protein; phosphoprotein; sodium channel; voltage-gated

Query Match 56.2%; Score 40.5; DB 2; Length 1682;
Best Local Similarity 54.5%; Pred. No. 3.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10
|||:|||||
Db 686 CGEWITWDC 696

Search completed: January 3, 2005, 15:52:58
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 15:30:50 ; Search time 188 Seconds
(without alignments)
30.605 Million cell updates/sec

Title: SEQ33
Perfect score: 72
Sequence: 1 cgywltiwc 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	63.9	218	1 Y232 SYNY3	Q55705 synchocyst
2	45	62.5	828	1 Q82BL4	Q82BL4 versinia pe
3	45	62.5	853	2 Q8DIA7	Q8DIA7 versinia pe
4	45	62.5	853	2 AAS60566	AAS60566 versinia
5	43.5	60.4	1933	2 P90870	P90870 alypsia cal
6	43	59.7	168	2 Q6H022	Q6H022 fremyella d
7	43	59.7	221	2 Q74JK6	Q74JK6 lactobacilli
8	43	59.7	221	2 AAS08923	AAS08923 lactobaci
9	43	59.7	501	2 Q6C574	Q6C574 yarrowia li
10	43	59.7	646	2 Q8DKD6	Q8DKD6 synchococc
11	43	59.7	733	2 Q8H2N7	Q8H2N7 oryza sativ
12	43	59.7	1084	2 Q7WY20	Q7WY20 pseudomonas
13	42.5	59.0	152	2 Q96AC0	Q96AC0 homo sapien
14	42.5	59.0	159	2 Q6RW13	Q6RW13 homo sapien
15	42.5	59.0	159	2 Q9NRW9	Q9NRW9 homo sapien
16	42.5	59.0	159	2 Q96PL4	Q96PL4 homo sapien
17	42.5	59.0	159	2 AAR25556	AAR25556 homo sapi
18	42.5	59.0	263	2 Q7YIF1	Q7YIF1 cryptospori
19	42.5	59.0	1765	2 Q8H457	Q8H457 rattus norv
20	42	58.3	72	1 VG35 BPMU5	Q05245 mycobacteri
21	42	58.3	133	2 Q855I3	Q855I3 mycobacteri
22	42	58.3	284	2 Q6MC01	Q6MC01 parachlamyd
23	42	58.3	284	2 CAF23898	CAF23898 parachlam
24	42	58.3	343	2 Q94FS2	Q94FS2 cajanus caj
25	42	58.3	389	2 Q83I51	Q83I51 tropheryma
26	42	58.3	441	2 Q922T6	Q922T6 rhizobium m
27	42	58.3	452	2 Q83G14	Q83G14 tropheryma
28	42	58.3	472	1 RBL_NITVU	Q59613 nitrobacter
29	42	58.3	473	2 Q9XD76	Q9XD76 nitrobacter
30	42	58.3	473	2 Q9XD77	Q9XD77 nitrobacter
31	42	58.3	473	2 Q8VQ84	Q8VQ84 nitrosospir

32 42 58.3 1308 2 Q8T6H2
33 42 58.3 1498 1 MRP6 MOUSE
34 42 58.3 1502 1 MRP6 RAT
35 42 58.3 1503 1 MRP6 HUMAN
36 41.5 57.6 122 2 Q8N1Z5
37 41.5 57.6 620 2 Q67544
38 41 56.9 266 2 Q82285
39 41 56.9 307 2 Q9M030
40 41 56.9 376 2 Q6CBE4
41 41 56.9 398 2 Q8IUD8
42 41 56.9 413 2 Q9HKA9
43 41 56.9 419 2 Q97VB7
44 41 56.9 432 2 Q6NUR0
45 41 56.9 432 2 AAH59368

Q8T6H2 dictyosteli
Q9r187 mus musculus
Q88269 rattus norv
Q95255 homo sapien
Q8N1Z5 homo sapien
Q67544 aquifex aeo
Q82285 enterococcu
Q9M030 arabidopsis
Q6CBE4 yarrowia li
Q8IUD8 homo sapien
Q9HKA9 thermoplas
Q97VB7 sulfolobus
Q6NUR0 homo sapien
AAH59368 homo sapi

ALIGNMENTS

RESULT 1
Y232 SYNY3
ID Y232 SYNY3 STANDARD; PRT; 218 AA.
AC Q55705;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical protein slr0232.
GN OrderedLocustNames=slr0232;
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96327529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res 2:153-166(1995).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the deda family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D64000; BAA10237.1; -;
CC FIR; S76385; S76385.
CC InterPro; IPR000252; Deda.
CC Pfam; PF00597; Deda; 1.
CC Complete proteome; Hypothetical protein.
SQ SEQUENCE 218 AA; 23781 MW; C04B5D7B7EA7F863 CRC64;

Query Match 63.9%; Score 46; DB 1; Length 218;
Best Local Similarity 66.7%; Pred. No. 50;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGVWLTIMG 9
Db 73 CGVWVGWRWG 81

RESULT 2
Q8ZBL4 PRELIMINARY; PRT; 828 AA.
ID Q8ZBL4
AC Q8ZBL4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

```
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ATP-dependent helicase.
GN Name=hrpB; OrderedLocusNames=YPO33394;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11596360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks R.M., Cerdano-Parraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.,
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL; AJ414156; CAC92624.1; -.
DR PIR; AD0412; AD0412.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR010225; DEAD_box_HrpB.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR007502; Helicase_dom.
DR Pfam; PF002070; DEAD; 1.
DR Pfam; PF04408; HA2; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR PIRSF; PIRSF005496; ATP_hel_hrpB; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR TIGRFAMs; TIGR01970; DEAH_box_HrpB; 1.
KW ATP-binding; Complete proteome; Helicase; Hydrolase.
SQ SEQUENCE 828 AA; 91982 MW; 3A1F883D5D5E583 CRC64;

Query Match 62.5%; Score 45; DB 2; Length 828;
Best Local Similarity 58.3%; Pred. No. 2.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 1 CGYWLTI--WGC 10
||| : |||
Db 372 CGLWLELLSWG 383

RESULT 3
ID Q8D1A7 PRELIMINARY; PRT; 853 AA.
AC Q8D1A7; Q74XU2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Helicase, ATP-dependent.
GN Name=hrpB; OrderedLocusNames=YPO291, y0794;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";

J. Bacteriol. 184:4601-4611(2002).
[2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RX Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS013681; AAM84381.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0004386; F:helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR010225; DEAD_box_HrpB.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR007502; Helicase_dom.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF04408; HA2; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR PIRSF; PIRSF005496; ATP_hel_hrpB; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR TIGRFAMs; TIGR01970; DEAH_box_HrpB; 1.
KW ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 853 AA; 95004 MW; B3DB738A18665B42 CRC64;

Query Match 62.5%; Score 45; DB 2; Length 853;
Best Local Similarity 58.3%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 1 CGYWLTI--WGC 10
||| : |||
Db 397 CGLWLELLSWG 408

RESULT 4
ID AAS60566 PRELIMINARY; PRT; 853 AA.
AC AAS60566;
DT 24-MAR-2004 (TrEMBLrel. 27, Created)
DT 24-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 04-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Helicase, ATP-dependent.
GN HRPB OR YPO291.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RX Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS017128; AAS60566.1; -.
KW Helicase.
SQ SEQUENCE 853 AA; 95004 MW; B3DB738A18665B42 CRC64;

Query Match 62.5%; Score 45; DB 2; Length 853;
Best Local Similarity 58.3%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 1 CGYWLTI--WGC 10
||| : |||
Db 397 CGLWLELLSWG 408
```



```

RESULT 5
P90670
ID P90670 PRELIMINARY; PRT; 1993 AA.
AC P90670;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sodium channel alpha-subunit SCAP1.
OS Aplysia californica (California sea hare).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Heterobranchia; Euthyneura; Opisthobranchia; Anaspiidea;
OC Aplysioidae; Aplysiidae; Aplysia.
OX NCBI_TaxID=6500;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97238630; PubMed=9115644;
RX Dyer J.R., Johnston W.L., Castellucci V.F., Dunn R.J.;
RT "Cloning and tissue distribution of the Aplysia Na+ channel alpha-
subunit cDNA.";
RL DNA Cell Biol. 16:347-356(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sodium channel family.
DR EMBL; U66915; AAC47457.1; -.
DR PIR; T30902; T30902.
DR HSP; P04775; IBYV.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:000518; C:voltage-gated sodium channel complex; IEA.
DR GO; GO:0005261; F:cation channel activity; IEA.
DR GO; GO:0005248; F:voltage-gated sodium channel activity; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001682; Ca/Na pore.
DR InterPro; IPR002111; Cat_channel_TrpL.
DR InterPro; IPR005821; Ion_trans.
DR InterPro; IPR005820; M_channel_nlg.
DR InterPro; IPR001696; Na_channel.
DR InterPro; IPR010526; Na_trans_assoc.
DR Pfam; PF00520; Ion_trans; 4.
DR Pfam; PF06512; Na_trans_assoc; 1.
DR PRINTS; PR00170; NACHANNEL.
DR KX Ion transport; Ionic channel; Sodium channel; Transmembrane;
KW Transport; Voltage-gated channel.
SQ SEQUENCE 1993 AA; 225896 MW; 33E174B9BF07E1A7 CRC64;

Query Match 60.4%; Score 43.5; DB 2; Length 1993;
Best Local Similarity 54.5%; Pred. No. 8.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 CGYWL-TWGC 10
Db 942 CGEWIESMMGC 952

RESULT 6
Q6H022
ID Q6H022 PRELIMINARY; PRT; 168 AA.
AC Q6H022;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Fremyella diplosiphon (Calothrix PCC 7601).
OC Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Fremyella.
OX NCBI_TaxID=1197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PD33;
RA Stowe-Evans E.L., Ford J., Kehoe D.M.;
RT "Genomic DNA Microarray Analysis: Identification of New Genes
Regulated by Light Color in the Cyanobacterium Fremyella
diplosiphon.";
```

```

J. Bacteriol. 186:4338-4349(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PD33;
RA Stowe-Evans E., Ford J., Kehoe D.M.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY548455; AAT41947.1; -.
KW Hypothetical protein.
SQ SEQUENCE 168 AA; 18932 MW; 1C9DB963D5210332 CRC64;

Query Match 59.7%; Score 43; DB 2; Length 168;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGYWLTIWG 9
Db 122 CGYWLSLRG 130

RESULT 7
Q74JK6
ID Q74JK6 PRELIMINARY; PRT; 221 AA.
AC Q74JK6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hemolysin-like protein.
GN OrderedLocustNames=LJ1101;
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX Fridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwahlen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollet B., Mercenier A., Klaenhammer T., Arigoni F., Scheil M.A.;
RT "The genome sequence of the probiotic intestinal bacterium
Lactobacillus johnsonii NCC 533.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
DR EMBL; AB017203; AAS08923.1; -.
DR InterPro; IPR004254; HlyVIII-related.
DR InterPro; IPR005744; HlyIII.
DR Pfam; PF03006; HlyIII; 1.
DR TIGRPFAM; TIGR01085; hlyIII; 1.
KW Complete proteome.
SQ SEQUENCE 221 AA; 24721 MW; 1DBC78FF9810E152 CRC64;

Query Match 59.7%; Score 43; DB 2; Length 221;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYWLTIWG 9
Db 172 GFWLLVWG 179

RESULT 8
AAS08923
ID AAS08923 PRELIMINARY; PRT; 221 AA.
AC AAS08923;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hemolysin-like protein.
GN LJ1101.
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]
```

```

RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Pridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwahlen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollet B., Mercenier A., Klaenhammer T., Arigoni F., Schell M.A.;
RT "The genome sequence of the probiotic intestinal bacterium
RT Lactobacillus Johnsonii NCC 533";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
DR EMBL; AE017203; AAS08923.1; -.
SQ SEQUENCE 221 AA; 24721 MW; 1DBC78FF9810E152 CRC64;

Query Match 59.7%; Score 43; DB 2; Length 221;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GYWLTIWG 9
Db 172 GFWLLWVG 179

RESULT 9
Q6C574 PRELIMINARY; PRT; 501 AA.
AC Q6C574;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Chromosome E of strain CLIB99 of Yarrowia lipolytica.
GN ORFNames=VALI0E20471g;
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG GENOLEVURES;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Franchin L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boistrame A., Boyer J.J., Cattolico L., Confaniolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382131; CAG79783.1; -.
SQ SEQUENCE 501 AA; 56411 MW; 91F08EFA63C60FB CRC64;

Query Match 59.7%; Score 43; DB 2; Length 501;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GYWLTIWG 10
Db 157 GYLLTWYC 165

RESULT 10
Q8DKD6 PRELIMINARY; PRT; 646 AA.
AC Q8DKD6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tl10923 protein.
GN OrderedLocusNames=tl10923;
OS Synecococcus elongatus (Thermosynecococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=32046;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynecococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005372; BAC08475.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2CC; 1.
DR SMART; SM00331; PP2C_SIG; 1.
KW Complete proteome.
SQ SEQUENCE 646 AA; 71071 MW; 56C5A74F76652D56 CRC64;

Query Match 59.7%; Score 43; DB 2; Length 646;
Best Local Similarity 62.5%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 YWLTIWGC 10
Db 415 YWINRWGC 422

RESULT 11
Q8H2N7 PRELIMINARY; PRT; 733 AA.
AC Q8H2N7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein OJ1138_B05.118.
GN Names=OJ1138_B05.118;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Embryophyta; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1];
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005486; BAC16197.1; -.
DR Gramene; Q8H2N7; -.
DR InterPro; IPR007658; DUF594.
DR Pfam; PF04578; DUF594; 1.
KW Hypothetical protein
SQ SEQUENCE 733 AA; 82939 MW; E95884DAD1DC2AC9 CRC64;

Query Match 59.7%; Score 43; DB 2; Length 733;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 YWLTIWG 9
Db 328 YWTTWG 334

```

```

RESULT 12
Q7WY20 Q7WY20 PRELIMINARY; PRT; 1084 AA.
AC
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE RSCC.
GN Names=rcsc; ORFNames=RL038;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PA14;
RX PubMed=14983043;
RA He J., Baldini R.L., Deziel E., Saucier M., Zhang Q., Liberati N.T.,
RA Lee D., Urbach J., Goodman H.M., Rahme L.G.;
RT "The broad host range pathogen Pseudomonas aeruginosa strain PA14
RT carries two pathogenicity islands harboring plant and animal virulence
RT genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2530-2535(2004).
CC -1- SIMILARITY: Contains 1 histidine kinase domain.
DR EMBL; AY273869; AAP84165.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . . ; IEA.
DR InterPro; IPR003594; ATPbind ATPase.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinase_N.
DR InterPro; IPR008207; Hpt.
DR InterPro; IPR000014; PAS.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HisKA; 1.
DR Pfam; PF01627; Hpt; 1.
DR Pfam; PF00072; Response_reg; 1.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HisKA; 1.
DR SMART; SM00091; PAS; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PSS0109; HIS_KIN; 1.
DR PROSITE; PSS0894; HPT; 1.
DR PROSITE; PSS0110; RESPONSE_REGULATORY; 1.
KW Kinase; phosphorylation; sensory transduction; Transferase.
SQ SEQUENCE 1084 AA; 119129 MW; C953PDD2F273BFB CRC64;

Query Match 59.7%; Score 43; DB 2; Length 1084;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GYWLTWGC 10
DB 750 GAWLKAWGC 758

RESULT 13
Q96AC0 Q96AC0 PRELIMINARY; PRT; 152 AA.
AC
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGTRAP protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA Rajkumar N., Yi Q., Nickerson D.A.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY488088; AAR25556.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR009436; AGTRAP.
DR Pfam; PF06396; AGTRAP; 1.
SQ SEQUENCE 159 AA; 17419 MW; 7E1D5C7E79AE6BC5 CRC64;

Query Match 59.0%; Score 42.5; DB 2; Length 159;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 2 GYW-LTWGC 10
DB 14 GHWLLTWGC 23

RESULT 14
Q6RW13 Q6RW13 PRELIMINARY; PRT; 159 AA.
AC
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Angiotensin II receptor-associated protein.
GN Name=AGTRAP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA Rajkumar N., Yi Q., Nickerson D.A.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY488088; AAR25556.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR009436; AGTRAP.
DR Pfam; PF06396; AGTRAP; 1.
SQ SEQUENCE 152 AA; 16669 MW; 637C01214175C3C9 CRC64;

Query Match 59.0%; Score 42.5; DB 2; Length 152;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 2 GYW-LTWGC 10
DB 14 GHWLLTWGC 23

```

Qy 2 GYW-LTIWGC 10
|:| || |||
Db 14 GHWLLTTWGC 23

RESULT 15

Q9NRW9 PRELIMINARY; PRT; 159 AA.
AC Q9NRW9; 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ATRAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Ye R.D., He R.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF165187; AAF89547.1; -.
DR Genew; HGNC:13539; ATRAP.
DR InterPro; IPR009436; ATRAP.
DR Pfam; PF06396; ATRAP; 1.
SQ SEQUENCE 159 AA; 17518 MW; 7E012C7E79AE6BC5 CRC64;

Query Match 59.0%; Score 42.5; DB 2; Length 159;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 2 GYW-LTIWGC 10
|:| || |||
Db 14 GHWLLTTWGC 23

Search completed: January 3, 2005, 15:52:15
Job time : 191 secs

Result No.	Score	Query Match	Length	DB	ID	Description	
1	72	100.0	10	5	ABP53932	Abp53932	VEGFR-3 b
2	54	75.0	10	5	ABP53931	Abp53931	VEGFR-3 b
3	50	69.4	10	5	ABP53968	Abp53968	VEGFR-3 b
4	39.5	54.9	9	5	ABJ04472	Abj04472	Stem cell
5	35.5	49.3	9	5	ABJ04461	Abj04461	Stem cell
6	35	48.6	9	5	ABJ04460	Abj04460	Stem cell
7	33.5	46.5	7	5	ABJ04531	Abj04531	Molt-4 le
8	33	45.8	7	5	ABP53964	Abp53964	VEGFR-3 b
9	33	45.8	8	2	AAV03715	AAV03715	Fluorine-
10	33	45.8	8	3	AAV76817	AAV76817	Immunogen
11	33	45.8	8	5	ABP53965	Abp53965	VEGFR-3 b
12	33	45.8	8	7	ADG94005	Adg94005	Immunogen
13	33	45.8	8	8	ADL98014	Adl98014	Peptide h
14	33	45.8	10	5	ABM46346	Abm46346	Desmoglei
15	33	45.8	10	5	ABM46607	Abm46607	Desmocoll
16	32	44.4	7	3	AAV76794	AAV76794	Somatosta
17	32	44.4	7	5	ABP53418	Abp53418	Backbone
18	32	44.4	9	4	ABP22609	Abp22609	HIV A11 m
19	32	44.4	9	4	ABP20314	Abp20314	HIV A03 m
20	32	44.4	9	5	AAU90543	Aau90543	Insulin/i
21	32	44.4	9	5	AAU90464	Aau90464	Insulin/i
22	32	44.4	9	5	AAU90542	Aau90542	Insulin/i
23	32	44.4	9	5	AAU90541	Aau90541	Insulin/i
24	32	44.4	10	2	AAW43886	AAW43886	Specific
25	32	44.4	10	3	AAV66209	AAV66209	HLA-A3-b1

CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 CC
 XX SQ Sequence 10 AA;

Query Match 100.0%; Score 72; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGYWLTIWGC 10
 |||||
 DB 1 CGYWLTIWGC 10

RESULT 2
 ABP53931
 ID ABP53931 standard; peptide; 10 AA.

XX AC ABP53931;

XX DT 09-JAN-2003 (first entry)

XX DE VEGFR-3 binding peptide SEQ ID NO:34.

XX KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
 KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
 KW cytosolic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
 KW vulnary; cell surface receptor; cancer; neovascularisation;
 KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
 KW diabetes; PDGF; platelet derived growth factor.

XX OS Homo sapiens.
 OS Synthetic.

XX FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "any amino acid"
 FT Misc-difference 10 /note= "any amino acid"
 FT Misc-difference 10 /note= "any amino acid"

XX PN WO200257299-A2.

XX PD 25-JUL-2002.

XX PF 16-JAN-2002; 2002WO-IB000099.

XX PR 17-JAN-2001; 2001US-0262476P.

XX PA (LUDW-) LUDWIG INST CANCER RES.
 XX PA (LICN) LICENTIA LTD.

XX PI Alitalo K, Koivunen E, Kubo H;

XX DR WPI; 2002-691521/74.

XX PT New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
 PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
 PT such as cancer and diseases of neovascularization.

XX PS Claim 12; Page 80; 149pp; English.

XX CC The present invention describes an isolated peptide (I) that binds to and
 CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnary activities, and can be used in gene therapy.
 CC Compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 CC
 XX SQ Sequence 10 AA;

Query Match 75.0%; Score 54; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.67;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GYWLTIWG 9
 |||||
 DB 2 GYWLTIWG 9

RESULT 3
 ABP53968
 ID ABP53968 standard; peptide; 10 AA.

XX AC ABP53968;

XX DT 09-JAN-2003 (first entry)

XX DE VEGFR-3 binding peptide SEQ ID NO:73.

XX KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
 KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
 KW cytosolic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
 KW vulnary; cell surface receptor; cancer; neovascularisation;
 KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
 KW diabetes; PDGF; platelet derived growth factor.

XX OS Homo sapiens.
 OS Synthetic.

XX FH Key Location/Qualifiers
 FT Misc-difference 5.7 /note= "X is any amino acid"
 FT Misc-difference 9 /note= "X is any amino acid"
 FT Misc-difference 9 /note= "X is any amino acid"

XX PN WO200257299-A2.

XX PD 25-JUL-2002.

XX PF 16-JAN-2002; 2002WO-IB000099.

XX PR 17-JAN-2001; 2001US-0262476P.

XX PA (LUDW-) LUDWIG INST CANCER RES.
 XX PA (LICN) LICENTIA LTD.

XX PI Alitalo K, Koivunen E, Kubo H;

XX DR WPI; 2002-691521/74.

XX PT New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
 PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
 PT such as cancer and diseases of neovascularization.

XX PS Disclosure; Page 147; 149pp; English.

XX CC The present invention describes an isolated peptide (I) that binds to and
 CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnary activities, and can be used in gene therapy.
 CC Compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,
CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
CC chronic hepatitis, haemangiomas and diabetes. The present sequence
CC represents a VSGFR-3 binding peptide, which is given in the
XX exemplification of the present invention
SQ Sequence 10 AA;

Query Match 69.4%; Score 50; DB 5; Length 10;
Best Local Similarity 60.0%; Pred. No. 2.3;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGYWLTIWGC 10
||| |
Db 1 CGYXXXXWC 10

RESULT 4
ABJ04472
ID ABJ04472 standard; peptide; 9 AA.
XX
AC ABJ04472;
XX
DT 24-OCT-2002 (first entry)
XX
DE Stem cell (mesenchymal) targeting peptide 61.
XX
KW BRASIL; targeting peptide; bacterial infection;
XX Biopanning and Rapid Analysis of Selective Interactive Ligands; diabetes;
KW inflammatory arthritis; atherosclerosis; cancer; autoimmune disease;
KW viral infection; cardiovascular disease; degenerative disease.
XX
OS Unidentified.
XX
PN WO200220822-A2.
XX
PD 14-MAR-2002.
XX
PF 07-SEP-2001; 2001WO-US028124.
XX
PR 08-SEP-2000; 2000US-0231266P.
PR 17-JAN-2001; 2001US-00765101.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Arap W, Pasqualini R;
XX
DR WPI; 2002-404697/43.
XX
PT Identification of targeting peptides that can be used to treat diseases
PT e.g. cancer and arthritis, by the BRASIL (Biopanning and Rapid Analysis
PT of Selective Ligands) method comprises a single differential
PT centrifugation step.
XX
PS Example 5; Page 76; 167pp; English.
XX
CC The invention comprises a method (BRASIL - Biopanning and Rapid Analysis
CC of Selective Interactive Ligands) to obtain a targeting peptide. The
CC BRASIL method of the invention involves: exposing a target to a phage
CC display library in a first phase; exposing the first phase to a second
CC phase; and separating the phage bound to the target from unbound phage.
CC The BRASIL method of the invention allows cell phages to be separated
CC from the remaining unbound phage in a single differential centrifugation
CC step. When compared to conventional cell panning methods, the BRASIL
CC method shows a significant increase in recovery of specific phage and a
CC substantial decrease in background. The BRASIL method is useful for
CC identifying targeting peptides. The targeting peptides identified by the
CC method of the invention are useful for treating disease states, such as:
CC diabetes; inflammatory arthritis; atherosclerosis; cancer; autoimmune
CC disease; bacterial infection; viral infection; cardiovascular disease and
CC degenerative disease. The present amino acid sequence represents a

CC targeting peptide of the invention
XX Sequence 9 AA;
SQ

Query Match 54.9%; Score 39.5; DB 5; Length 9;
Best Local Similarity 60.0%; Pred. No. 1.7e+06;
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGYWLTIWGC 10
||| |
Db 1 CG-WFSWGC 9

RESULT 5
ABJ04461
ID ABJ04461 standard; peptide; 9 AA.
XX
AC ABJ04461;
XX
DT 24-OCT-2002 (first entry)
XX
DE Stem cell (mesenchymal) targeting peptide 50.
XX
KW BRASIL; targeting peptide; bacterial infection;
KW Biopanning and Rapid Analysis of Selective Interactive Ligands; diabetes;
KW inflammatory arthritis; atherosclerosis; cancer; autoimmune disease;
KW viral infection; cardiovascular disease; degenerative disease.
XX
OS Unidentified.
XX
PN WO200220822-A2.
XX
PD 14-MAR-2002.
XX
PF 07-SEP-2001; 2001WO-US028124.
XX
PR 08-SEP-2000; 2000US-0231266P.
PR 17-JAN-2001; 2001US-00765101.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Arap W, Pasqualini R;
XX
DR WPI; 2002-404697/43.
XX
PT Identification of targeting peptides that can be used to treat diseases
PT e.g. cancer and arthritis, by the BRASIL (Biopanning and Rapid Analysis
PT of Selective Ligands) method comprises a single differential
PT centrifugation step.
XX
PS Example 5; Page 76; 167pp; English.
XX
CC The invention comprises a method (BRASIL - Biopanning and Rapid Analysis
CC of Selective Interactive Ligands) to obtain a targeting peptide. The
CC BRASIL method of the invention involves: exposing a target to a phage
CC display library in a first phase; exposing the first phase to a second
CC phase; and separating the phage bound to the target from unbound phage.
CC The BRASIL method of the invention allows cell phages to be separated
CC from the remaining unbound phage in a single differential centrifugation
CC step. When compared to conventional cell panning methods, the BRASIL
CC method shows a significant increase in recovery of specific phage and a
CC substantial decrease in background. The BRASIL method is useful for
CC identifying targeting peptides. The targeting peptides identified by the
CC method of the invention are useful for treating disease states, such as:
CC diabetes; inflammatory arthritis; atherosclerosis; cancer; autoimmune
CC disease; bacterial infection; viral infection; cardiovascular disease and
CC degenerative disease. The present amino acid sequence represents a
XX targeting peptide of the invention
SQ Sequence 9 AA;

Query Match 49.3%; Score 35.5; DB 5; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.7e+06;

Matches 5; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGWLTWGC 10
 ||:|:|
 Db 1 CGWW-GLWPC 9

RESULT 6
 ABJ04460
 ID ABJ04460 standard; peptide; 9 AA.

XX AC ABJ04460;
 XX AC
 XX AC
 DT 24-OCT-2002 (first entry)
 XX DE
 XX DE Stem cell (mesenchymal) targeting peptide 49.
 XX XX
 KW BRASIL; targeting peptide; bacterial infection;
 KW Biopanning and Rapid Analysis of Selective Interactive Ligands; diabetes;
 KW inflammatory arthritis; atherosclerosis; cancer; autoimmune disease;
 KW viral infection; cardiovascular disease; degenerative disease.
 XX OS Unidentified.
 XX XX
 XX XX
 PN WO200220822-A2.
 XX XX
 PD 14-MAR-2002.
 XX XX
 XX 07-SEP-2001; 2001WO-US028124.
 XX XX
 XX 08-SEP-2000; 2000US-0231266P.
 PR 17-JAN-2001; 2001US-00765101.
 XX XX

PA (TEXA) UNIV TEXAS SYSTEM.
 XX XX
 XX Arap W, Pasqualini R;
 XX WPI; 2002-404697/43.
 DR XX
 XX XX

PT Identification of targeting peptides that can be used to treat diseases
 PT e.g. cancer and arthritis, by the BRASIL (Biopanning and Rapid Analysis
 PT of Selective Ligands) method comprises a single differential
 PT centrifugation step.

PS Example 5; Page 76; 167pp; English.

XX The invention comprises a method (BRASIL - Biopanning and Rapid Analysis
 CC of Selective Interactive Ligands) to obtain a targeting peptide. The
 CC BRASIL method of the invention involves: exposing a target to a phage
 CC display library in a first phase; exposing the first phase to a second
 CC phase; and separating the phage bound to the target from unbound phage.
 CC The BRASIL method of the invention allows cell phages to be separated
 CC from the remaining unbound phage in a single differential centrifugation
 CC step. When compared to conventional cell panning methods, the BRASIL
 CC method shows a significant increase in recovery of specific phage and a
 CC substantial decrease in background. The BRASIL method is useful for
 CC identifying targeting peptides. The targeting peptides identified by the
 CC method of the invention are useful for treating disease states, such as:
 CC diabetes; inflammatory arthritis; atherosclerosis; cancer; autoimmune
 CC disease; bacterial infection; viral infection; cardiovascular disease and
 CC degenerative disease. The present amino acid sequence represents a
 CC targeting peptide of the invention

XX Sequence 9 AA;

Query Match 48.6%; Score 35; DB 5; Length 9;
 Best Local Similarity 50.0%; Pred. No. 1.7e+06;
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGWLTW 8
 |:|:|
 Db 1 CDWTTAW 8

RESULT 7
 ABJ04531
 ID ABJ04531 standard; peptide; 7 AA.

XX AC ABJ04531;
 XX AC
 XX 24-OCT-2002 (first entry)
 XX XX
 DE Molt-4 leukaemia cell line targeting peptide 16.

XX BRASIL; targeting peptide; bacterial infection;
 KW Biopanning and Rapid Analysis of Selective Interactive Ligands; diabetes;
 KW inflammatory arthritis; atherosclerosis; cancer; autoimmune disease;
 KW viral infection; cardiovascular disease; degenerative disease.

XX OS Unidentified.

XX PN WO200220822-A2.

XX PD 14-MAR-2002.

XX PF 07-SEP-2001; 2001WO-US028124.

XX PR 08-SEP-2000; 2000US-0231266P.

XX PR 17-JAN-2001; 2001US-00765101.

XX PA (TEXA) UNIV TEXAS SYSTEM.

XX PI Arap W, Pasqualini R;

XX WPI; 2002-404697/43.

PT Identification of targeting peptides that can be used to treat diseases
 PT e.g. cancer and arthritis, by the BRASIL (Biopanning and Rapid Analysis
 PT of Selective Ligands) method comprises a single differential
 PT centrifugation step.

PS Claim 79; Page 100; 167pp; English.

XX The invention comprises a method (BRASIL - Biopanning and Rapid Analysis
 CC of Selective Interactive Ligands) to obtain a targeting peptide. The
 CC BRASIL method of the invention involves: exposing a target to a phage
 CC display library in a first phase; exposing the first phase to a second
 CC phase; and separating the phage bound to the target from unbound phage.
 CC The BRASIL method of the invention allows cell phages to be separated
 CC from the remaining unbound phage in a single differential centrifugation
 CC step. When compared to conventional cell panning methods, the BRASIL
 CC method shows a significant increase in recovery of specific phage and a
 CC substantial decrease in background. The BRASIL method is useful for
 CC identifying targeting peptides. The targeting peptides identified by the
 CC method of the invention are useful for treating disease states, such as:
 CC diabetes; inflammatory arthritis; atherosclerosis; cancer; autoimmune
 CC disease; bacterial infection; viral infection; cardiovascular disease and
 CC degenerative disease. The present amino acid sequence represents a
 CC targeting peptide of the invention

XX Sequence 7 AA;

Query Match 46.5%; Score 33.5; DB 5; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.7e+06;
 Matches 5; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

QY 1 CGWLTWGC 10
 |:|:|:|
 Db 1 CSWV---WGC 7

RESULT 8
 ABP53964
 ID ABP53964 standard; peptide; 7 AA.
 XX AC ABP53964;

XX 09-JAN-2003 (first entry)
XX VEGFR-3 binding peptide SEQ ID NO:67.
XX Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
XX angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
XX cytoskeletal; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
XX vulnary; cell surface receptor; cancer; neovascularisation;
XX liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
XX diabetes; PDGF; platelet derived growth factor.
XX Homo sapiens.
OS Synthetic.
OS Key Location/Qualifiers
FT Misc-difference 4. .6 /note= "X is any amino acid"
FT WO200257299-A2.
XX 25-JUL-2002.
XX 16-JAN-2002; 2002WO-IB000099.
XX 17-JAN-2001; 2001US-0262476P.
XX (LUDW-) LUDWIG INST CANCER RES.
PA (LICN) LICENTIA LTD.
XX Alitalo K, Koivunen E, Kubo H;
PI WPI; 2002-691521/74.
DR New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
XX diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
XX such as cancer and diseases of neovascularization.
XX Claim 21; Page 81; 149pp; English.
XX The present invention describes an isolated peptide (I) that binds to and
XX inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
XX have cytostatic, hepatotropic, antiinflammatory, hypotensive,
XX antidiabetic and vulnerary activities, and can be used in gene therapy.
XX Compositions and methods from the present invention are useful for
XX diagnosing, evaluating and treating disorders mediated by the activity of
XX the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
XX liver, spleen, kidney, lymph node, small intestine, blood cells,
XX pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
XX skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
XX neovascularisation, e.g. liver diseases, hypertension, post-trauma,
XX chronic hepatitis, haemangiomas and diabetes. The present sequence
XX represents a specifically claimed VEGFR-3 binding peptide from the
XX present invention
SQ Sequence 7 AA;
Query Match 45.8%; Score 33; DB 5; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.7e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 GYWLTIW 8
|||
Db 1 GYWXW 7
RESULT 9
AAV03715
ID AAV03715 standard; peptide; 8 AA.
XX AAV03715;
AC AAV03715;
XX 08-JUN-1999 (first entry)
XX Fluorine-18 (F-18) labeled peptide 2.
XX 18F radionuclide; targeting vector; positron emission tomography; F-18;
XX radiolabeling; thiol; fluorine-18.
XX Synthetic.
XX Key Location/Qualifiers
FT Misc-difference 1 /note= "N-terminal acetylation; optionally has a free or
FT protected thiol group"
FT Misc-difference 2 /note= "D-form residue"
FT Misc-difference 3 /note= "D-form residue"
FT Misc-difference 5 /note= "D-form residue; optionally has a free or
FT protected thiol group"
FT Misc-difference 7 /note= "D-form residue"
FT Misc-difference 8 /note= "D-form residue"
XX WO9911590-A1.
XX 11-MAR-1999.
XX 03-SEP-1998; 98WO-US018268.
XX 03-SEP-1997; 97US-0057485P.
XX (IMMU-) IMMUNOMEDICS INC.
XX Griffiths GL;
XX WPI; 1999-228967/19.
XX Radiolabeling thiol-containing peptides with fluorine-18.
XX Claim 14; Page 15; 22pp; English.
XX The invention relates to a method for incorporating 18F radionuclide into
XX peptide-containing targeting vectors for use in clinical positron
XX emission tomography. Radiolabeling thiol-containing peptides with
XX fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol
XX group with a labeling reagent of formula: 18F-(CH2)m-CH(R2)-(CH2)n-X, or a
XX fluorinated alkene in which at least one of the two double bonded carbon
XX atoms bears at least one leaving group comprising I, Br, Cl, azide,
XX tosylate, mesylate, nosylate or triflate. n, m = 0-2; n+m = 0-2; X = I,
XX Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide
XX (optionally substituted by 1-2 alkyl) or 3-sulfonamide; R1, R2 = I,
XX OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl (optionally
XX substituted by CONH2, COOH, OH, sulfonic acid, tertiary amine or
XX quaternary ammonium), COOR', CONR'2 or COR'; and R' = 1-6C alkyl or
XX phenyl. The method is used for radiolabeling peptide-containing targeting
XX vectors such as proteins, antibodies, antibody fragments and receptor-
XX targeted peptides for use in routine clinical positron emission
XX tomography. The method is simple and efficient. The method uses the
XX unique property of the free thiol groups which are rapidly alkylated at
XX neutral pH and moderate temperature. Sequences AAY03714-716 represent
XX examples of F-18 labeled peptides used in the method of detecting a
XX tissue
XX Sequence 8 AA;
Query Match 45.8%; Score 33; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGYW 4
|||

```

Db          5 CGYW 8

RESULT 10
AAV76817
ID AAV76817 standard; peptide; 8 AA.
XX
XX
AC AAV76817;
XX
XX
DT 28-APR-2000 (first entry)
XX
DE Immunogenic peptide for bi-specific antibody recognition.
XX
XX Immunogenic peptide; bi-specific antibody; diagnosis; immune response;
KW diseased tissue identification; therapy.
XX
XX Synthetic.
XX
XX
PH Key Location/Qualifiers
FT Misc-difference 1
   /note= "acetylated; modified with free amino acid group,
   protected amino acid group, chelating agent or a metal-
   chelate complex"
FT Misc-difference 2
   /note= "D-form residue"
FT Misc-difference 3
   /note= "D-form residue"
FT Misc-difference 5
   /note= "D-form residue"
FT Misc-difference 7
   /note= "D-form residue"
FT Misc-difference 8
   /note= "D-form residue"
FT
FT
XX WO9966951-A2.
XX
XX 29-DEC-1999.
XX
XX 22-JUN-1999; 99WO-US013879.
XX
XX 22-JUN-1998; 98US-0090142P.
XX 14-OCT-1998; 98US-0104156P.
XX
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;
XX WPI; 2000-160561/14.
XX
XX Bi-specific antibodies that bind specific target tissue and targeted
XX conjugates.
XX
XX Claim 22; Page 61; 76pp; English.
XX
XX This sequence represents an immunogenic peptide for a bi-specific
XX antibody. The invention relates to a method of treating or identifying
XX diseased tissues in a patient comprising administering a bi-specific
XX antibody (or fragment) having at at least 1 arm (A) that specifically
XX binds a targeted tissue and at least 1 arm (B) that specifically binds a
XX targetable conjugate. The methods and bi-specific antibodies and fusion
XX proteins are useful for pre-targeting methods of diagnosis and therapy.
XX It is advantageous to raise bi-specific antibodies against a targetable
XX conjugate that is capable of carrying at least 1 diagnostic or
XX therapeutic agent. The characteristics of the chelator, metal chelate
XX complex, therapeutic agent or diagnostic agent can be varied to
XX accommodate differing applications without raising new bi-specific
XX antibodies for each new application. The targetable conjugate is selected
XX to elicit sufficient immune responses and also for rapid in vivo
XX clearance when used within the bi-specific antibody targeting method
XX
XX Sequence 8 AA;

Query Match          45.8%; Score 33; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGYW 4
   ||||
Db 5 CGYW 8

RESULT 11
ABP53965
ID ABP53965 standard; peptide; 8 AA.
XX
XX ABP53965;
XX
DT 09-JAN-2003 (first entry)
XX
XX VEGFR-3 binding peptide SEQ ID NO:68.
XX
XX Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
XX angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
XX cytostatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
XX vulnary; cell surface receptor; cancer; neovascularisation;
XX liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
XX diabetes; PDGF; platelet derived growth factor.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 4. .6
   /note= "X is any amino acid"
FT Misc-difference 8
   /note= "any amino acid"
FT
FT
XX WO200257299-A2.
XX
XX 25-JUL-2002.
XX
XX 16-JAN-2002; 2002WO-1B000099.
XX
XX 17-JAN-2001; 2001US-0262476P.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX (LICN ) LICENTIA LTD.
XX
XX Alitalo K, Koivunen E, Kubo H;
XX WPI; 2002-691521/74.
XX
XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
XX diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
XX such as cancer and diseases of neovascularization.
XX
XX Claim 22; Page 81; 149pp; English.
XX
XX The present invention describes an isolated peptide (I) that binds to and
XX inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
XX have cytostatic, hepatotropic, antiinflammatory, hypotensive,
XX antidiabetic and vulnerary activities, and can be used in gene therapy.
XX Compositions and methods from the present invention are useful for
XX diagnosing, evaluating and treating disorders mediated by the activity of
XX the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
XX liver, spleen, kidney, lymph node, small intestine, blood cells,
XX pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
XX skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
XX neovascularisation, e.g. liver diseases, hypertension, post-trauma,
XX chronic hepatitis, haemangiomas and diabetes. The present sequence
XX represents a specifically claimed VEGFR-3 binding peptide from the
XX present invention
XX
XX Sequence 8 AA;

```

Query Match 45.8%; Score 33; DB 5; Length 8;
Best Local Similarity 57.1%; Pred. No. 1.7e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 GYWLTIW 8
| | | |
Db 1 GYWXXW 7

RESULT 12
ADG94005
ID ADG94005 standard; peptide; 8 AA.
XX AC ADG94005;
DT 11-MAR-2004 (first entry)
XX DE Immunogenic peptide.
XX KW Immunogenic peptide; multi-specific antibody; polymer conjugate; tumour;
KW cytostatic; photodynamic therapy.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 1
FT Misc-difference 2 /note= "Optionally methylated or Acetylated"
FT Misc-difference 3 /note= "D-form residue"
FT Misc-difference 3 /note= "D-form residue"
FT Misc-difference 5 /note= "D-form residue"
FT Misc-difference 7 /note= "Optionally methylated or Acetylated D-form residue"
FT Misc-difference 7 /note= "D-form residue"
FT Misc-difference 8 /note= "D-form residue"
FT Misc-difference 8 /note= "D-form residue"
XX US2003026764-A1.
XX PD 06-FEB-2003.
XX PF 31-JUL-2002; 2002US-00209592.
XX PR 31-JUL-2001; 2001US-0308605P.
XX PA (IMMU-) IMMUNOMEDICS INC.
XX PI Griffiths GL;
XX DR WPI; 2003-801085/75.
XX Targeting an agent towards a target site in a tissue, by administering in
PT a tissue, a multi-specific antibody or its fragment and a polymer
PT conjugate that binds to the capture arm of the multi-specific antibody.
XX Disclosure; Page 14; 19pp; English.
XX The invention relates to targeting an agent towards a target site in a
CC tissue, comprising administering to the tissue, a multi-specific antibody
CC or its fragment, comprising a targeting arm that binds to an antigen of
CC the target site and a capture arm that binds to a polymer conjugate, and
CC administering a polymer conjugate that binds to the capture arm, the
CC conjugate has a multi-specific antibody or its fragment, and a therapeutic
CC agent, a peptide, an enzyme and a labelled ligand. Also included is a kit
CC useful for targeting a target site within a tissue in a subject or tissue
CC sample comprising the above mentioned multi-specific antibody or its
CC fragment and a polymer conjugate. The method is used for targeting an
CC agent towards a target site in a tissue (e.g. a tumour). The method is
CC also useful for therapeutic or diagnostic purposes and further in
CC photodynamic therapy. The present sequence is an immunogenic peptide used

CC in the method of the invention.
XX Sequence 8 AA;
SQ Query Match 45.8%; Score 33; DB 7; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGYW 4
| | | |
Db 5 CGYW 8

RESULT 13
ADL98014
ID ADL98014 standard; peptide; 8 AA.
XX AC ADL98014;
DT 20-MAY-2004 (first entry)
XX DE Peptide hapten #2.
XX KW photodynamic diagnosis; cancer; tumour; cardiovascular lesion;
KW inflammatory disease; neurodegenerative disease; metabolic disease;
KW infectious disease; B-cell malignancy; Alzheimer's disease; amyloidosis;
KW autoimmune disease; bacterial infection; fungal infection;
KW parasitic infection; viral infection;
KW carcinoembryonic antigen-expressing tumour.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 1 /note= "Optionally Cysteiny1; if Cys the residue is
FT methylated; N-terminal acetylated."
FT Misc-difference 2 /note= "D form residue"
FT Misc-difference 3 /note= "D form residue"
FT Modified-site 5 /note= "Optionally Cysteiny1; if Cys the residue is
FT methylated; N-terminal acetylated."
FT Misc-difference 7 /note= "D form residue"
FT Misc-difference 8 /note= "D form residue"
FT Misc-difference 8 /note= "D form residue"
XX US2004043030-A1.
XX PN 04-MAR-2004.
XX PD 09-JUN-2003; 2003US-00456580.
XX PF 31-JUL-2001; 2001US-0308605P.
XX PR 31-JUL-2002; 2002US-00209592.
XX PA (IMMU-) IMMUNOMEDICS INC.
XX PI Griffiths GL, Goldenberg DM, Hansen HJ;
XX DR WPI; 2004-313738/29.
XX Treating cancer and metabolic diseases by administering a multi-specific
PT antibody having a targeting arm that binds to an antigen and a capture
PT arm that binds to a polymer conjugate comprising a therapeutic agent.
XX Disclosure; Page 16; 24pp; English.
XX The invention relates to a method of diagnosing or treating a disease or
CC disorder. The method involves administering to a tissue a multi-specific
CC antibody (I) or antibody fragment, comprising a targeting arm that binds
CC to an antigen on the target site, and a capture arm that binds to a

polymer conjugate, and administering to the tissue a polymer conjugate that binds to the capture arm, the polymer conjugate comprising a polymer conjugated to a diagnostic or therapeutic agent. Also included is a method for photodynamic diagnosis or treatment of a disease or disorder; or intravascular or endoscopic method for diagnosing or treating a disease or disorder. The method is useful for diagnosing or treating a disease or disorder chosen from cancer (oesophageal, gastric, colonic, rectal, pancreatic, lung, breast, ovarian, urinary bladder, endometrial, cervical, testicular, renal, adrenal and liver cancer, solid tumour, B-cell malignancy or T-cell malignancy); cardiovascular lesion; an inflammatory disease; neurodegenerative disease; metabolic disease; and an infectious disease. The B-cell malignancy is chosen from indolent forms of B-cell lymphomas, aggressive forms of B-cell lymphomas, chronic lymphatic leukaemias, acute lymphatic leukaemias, and multiple myeloma. The solid tumour is chosen melanoma, carcinoma (preferably renal carcinoma, lung carcinoma, intestinal carcinoma, and stomach carcinoma), glioma and sarcoma. The cardiovascular lesion is chosen from infarct, clot, embolus, atherosclerotic plaque and ischaemia. The neurodegenerative disease is Alzheimer's disease. The metabolic disease is amyloidosis, where the antibody binds amyloid. The disease or disorder is displaced or ectopic normal tissue chosen from endometrium, thymus, spleen and parathyroid. The method can be used for normal tissue ablation, where the tissue is chosen from bone marrow and spleen. The disease or disorder is an autoimmune disease such as myasthenia gravis, lupus nephritis, lupus erythematosus, and rheumatoid arthritis. Class III autoimmune diseases such as immune-mediated thrombocytopenias, such as acute idiopathic thrombocytopenic purpura and chronic idiopathic thrombocytopenic purpura, dermatomyositis, Sjogren's syndrome, multiple sclerosis, Sydenham's chorea, myasthenia gravis, systemic lupus erythematosus, lupus nephritis, rheumatic fever, polyglandular syndromes, bullous pemphigoid, diabetes mellitus, Henoch-Schönlein purpura, post-streptococcal nephritis, erythema nodosum, Takayasu's arteritis, Addison's disease, rheumatoid arthritis, sarcoidosis, ulcerative colitis, erythema multiforme, IgA nephropathy, polyarthritis nodosa, ankylosing spondylitis, Goodpasture's syndrome, thromboangiitis obliterans, primary biliary cirrhosis, Hashimoto's thyroiditis, thyrotoxicosis, scleroderma, chronic active hepatitis, polymyositis/dermatomyositis, polychondritis, pemphigus vulgaris, Wegener's granulomatosis, membranous nephropathy, amyotrophic lateral sclerosis, tabes dorsalis, giant cell arteritis/polyalgia, pernicious anaemia, rapidly progressive glomerulonephritis, or fibrosing alveolitis. The infectious disease is chosen from bacterial, fungal, parasitic and viral lesion. The infectious disease is caused by a fungus chosen from Microsporium, Trichophyton, Epidermophyton, Sporothrix schenckii, Cryptococcus neoformans, Coccidioides immitis, Histoplasma capsulatum, Blastomyces dermatitidis, and Candida albicans. The infectious disease is caused by a virus chosen from HIV, herpes virus, cytomegalovirus, rabies virus, influenza virus, hepatitis B virus, Sendai virus, feline leukaemia virus, Reo virus, polio virus, human serum parvo-like virus, simian virus 40, respiratory syncytial virus, mouse mammary tumour virus, Varicella-Zoster virus, Dengue virus, rubella virus, measles virus, adenovirus, human T-cell leukaemia viruses, Epstein-Barr virus, murine leukaemia virus, mumps virus, vesicular stomatitis virus, Sindbis virus, lymphocytic choriomeningitis virus, wart virus and blue tongue virus. The infectious disease is caused by a bacterium chosen from Bacillus anthracis, Streptococcus agalactiae, Legionella pneumophila, Streptococcus pyogenes, Escherichia coli, Neisseria gonorrhoeae, Neisseria meningitidis, Pneumococcus, Haemophilus influenzae B, Treponema pallidum, Lyme disease spirochetes, Pseudomonas aeruginosa, Mycobacterium leprae, Brucella abortus, Mycobacterium tuberculosis, and Tetanus toxin. The infectious disease is caused by a protozoa chosen from Plasmodium falciparum, Plasmodium vivax, Toxoplasma gondii, Trypanosoma rangeli, Trypanosoma cruzi, Trypanosoma rhodesiense, Trypanosoma brucei, Schistosoma mansoni, Schistosoma japonicum, Babesia bovis, Elmeria tenella, Onchocerca volvulus, Leishmania tropica, Trichinella spiralis, Onchocerca volvulus, Theileria parva, Taenia hydatigena, Taenia ovis, Taenia saginata, Echinococcus granulosus, and Mesocostoides corti. The infectious disease is caused by a mycoplasma chosen from Mycoplasma arthritis, M. hyorhinis, M. orale, M. arginini, Achlepleasma laidlawii, M. salivarium and M. pneumoniae. The cancer is preferably chosen from carcinoembryonic antigen (CEA)-expressing tumour or a CD20-expressing malignancy. The present sequence represents a peptide used in the method of the invention.

SQ Sequence 8 AA;

Query Match 45.8%; Score 33; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGYW 4
DB 5 CGYW 8

RESULT 14

ABB46346
ID ABB46346 standard; peptide; 10 AA.

XX ABB46346;

DT 30-JAN-2002 (first entry)

DE Desmoglein-2 CAR sequence cyclic peptide SEQ ID NO 1090.

KW Desmosomal cadherin; cell adhesion; CAR sequence; immunosuppressive;
KW cytostatic; antiapoptotic; wound healing; reduce scar tissue; skin graft;
KW organ implant; autoimmune blistering disorder; cancer; apoptosis; cyclic.
OS Synthetic.

XX WO200172956-A2.

PD 04-OCT-2001.

PF 27-MAR-2001; 2001WO-IB001400.

XX 27-MAR-2000; 2000US-00535852.

PA (ADHE-) ADHEREX TECHNOLOGIES INC.

PI Blaschuk OW, Symonds JM, Gour BJ;

DR WPI; 2002-025778/03.

XX Modulating agents for inhibiting or enhancing desmosomal cadherin mediated cell adhesion, useful for facilitating wound healing and/or reducing scar tissue, treating cancer and inducing apoptosis.

PS Claim 18; Page 101; 127pp; English.

XX The invention relates to modulating agents for inhibiting or enhancing desmosomal cadherin mediated cell adhesion, comprising a modulating agent comprising a desmosomal cadherin cell adhesion recognition CAR sequence (ABB45341-ABB47262), a non-peptide mimetic of a desmosomal cadherin CAR sequence, a substance such as an antibody or antigen-binding fragment that specifically binds a desmosomal cadherin CAR sequence and/or a polynucleotide encoding a polypeptide that comprises a desmosomal cadherin CAR sequence or analogue. The modulating agents have immunosuppressive, cytostatic and antiapoptotic activity and are used to facilitate wound healing and/or reduce scar tissue, for enhancing adhesion of foreign tissue implants (e.g. skin graft or organ implant), treating an autoimmune blistering disorder and to treat cancer (e.g. carcinoma, leukaemia or melanoma) and induce apoptosis

SQ Sequence 10 AA;

Query Match 45.8%; Score 33; DB 5; Length 10;
Best Local Similarity 60.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGYWLTIWGC 10

DB 1 CGYALDARGC 10

RESULT 15

ABB46607
ID ABB46607 standard; peptide; 10 AA.
XX
AC ABB46607;
XX
DT 30-JAN-2002 (first entry)
XX
DE Desmocollin-1 CAR cyclic peptide 9.
XX
KW Desmosomal cadherin; cell adhesion; CAR sequence; immunosuppressive;
KW cytotatic; antiapoptotic; wound healing; reduce scar tissue; skin graft;
KW organ implant; autoimmune blistering disorder; cancer; apoptosis; cyclic.
XX
OS Synthetic.
XX
PN WO200172956-A2.
XX
PD 04-OCT-2001.
XX
PF 27-MAR-2001; 2001WO-1B001400.
XX
PR 27-MAR-2000; 2000US-00535852.
XX
PA (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI Blaschuk OW, Symonds JM, Gour BJ;
XX
DR WPI; 2002-025778/03.
XX
PT Modulating agents for inhibiting or enhancing desmosomal cadherin
PT mediated cell adhesion, useful for facilitating wound healing and/or
PT reducing scar tissue, treating cancer and inducing apoptosis.
XX
PS Claim 23; Page 109; 127pp; English.
XX
CC The invention relates to modulating agents for inhibiting or enhancing
CC desmosomal cadherin mediated cell adhesion, comprising a modulating agent
CC comprising a desmosomal cadherin cell adhesion recognition CAR sequence
CC (ABB45341-ABB47262), a non-peptide mimetic of a desmosomal cadherin CAR
CC sequence, a substance such as an antibody or antigen-binding fragment
CC that specifically binds a desmosomal cadherin CAR sequence and/or a
CC polynucleotide encoding a polypeptide that comprises a desmosomal
CC cadherin CAR sequence or analogue. The modulating agents have
CC immunosuppressive, cytostatic and antiapoptotic activity and are used to
CC facilitate wound healing and/or reduce scar tissue, for enhancing
CC adhesion of foreign tissue implants (e.g. skin graft or organ implant),
CC treating an autoimmune blistering disorder and to treat cancer (e.g.
CC carcinoma, leukaemia or melanoma) and induce apoptosis
XX
SQ Sequence 10 AA;

Query Match 45.8%; Score 33; DB 5; Length 10;
Best Local Similarity 60.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGYWLTIWGC 10
||| |
Db 1 CGYATTADGC 10

Search completed: January 3, 2005, 16:07:18
Job time : 153 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:02:29 ; Search time 37 Seconds
(without alignments)
17.924 Million cell updates/sec

Title: SEQ33

Perfect score: 72

Sequence: 1 cgywtiwcg 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 110780

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
 - 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
 - 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
 - 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
 - 5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
 - 6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	45.8	10	4	US-09-535-852-1090
2	33	45.8	10	4	US-09-535-852-1352
3	30	41.7	8	2	US-08-520-535-12
4	30	41.7	8	2	US-09-079-432-12
5	30	41.7	9	2	US-08-986-234-92
6	30	41.7	9	4	US-09-311-784A-374
7	30	41.7	9	4	US-09-790-497A-1
8	30	41.7	10	3	US-09-108-709-22
9	30	41.7	10	4	US-09-790-497A-110
10	29	40.3	5	1	US-07-946-237-4
11	29	40.3	5	2	US-08-530-566-10
12	29	40.3	5	3	US-09-195-726-10
13	29	40.3	5	3	US-09-067-755-10
14	29	40.3	5	4	US-08-239-765C-4
15	29	40.3	7	4	US-09-069-827A-94
16	29	40.3	9	4	US-09-311-784A-348
17	29	40.3	10	2	US-08-735-253-8
18	29	40.3	10	2	US-08-735-253-13
19	29	40.3	10	3	US-08-481-968A-21
20	29	40.3	10	3	US-08-154-712B-21
21	29	40.3	10	4	US-09-947-925A-21
22	28	38.9	8	3	US-09-315-304B-1649
23	28	38.9	10	1	US-08-250-789A-119
24	28	38.9	10	4	US-09-462-917A-73
25	28	38.9	10	4	US-09-125-641-3
26	28	38.9	10	4	US-09-790-497A-5
27	28	38.9	10	4	US-09-790-497A-24

28	27.5	38.2	6	1	US-08-191-571-12	Sequence 12, Appl
29	27.5	38.2	6	5	PCT-US95-00296-12	Sequence 12, Appl
30	27	37.5	8	3	US-08-925-002-12	Sequence 12, Appl
31	27	37.5	8	3	US-08-586-670A-17	Sequence 17, Appl
32	27	37.5	8	3	US-09-082-279B-1495	Sequence 1495, Ap
33	27	37.5	8	4	US-09-834-784-1495	Sequence 1495, Ap
34	27	37.5	8	4	US-09-910-552-12	Sequence 12, Appl
35	27	37.5	8	4	US-09-350-641C-1650	Sequence 1650, Ap
36	27	37.5	10	3	US-09-315-304B-1587	Sequence 1587, Ap
37	27	37.5	10	4	US-09-350-325-47	Sequence 47, Appl
38	27	37.5	10	4	US-09-535-852-1357	Sequence 1357, Ap
39	27	37.5	10	4	US-09-350-641C-1587	Sequence 1587, Ap
40	27	37.5	10	4	US-09-239-043D-2474	Sequence 2474, Ap
41	27	37.5	10	4	US-09-620-091-28	Sequence 28, Appl
42	27	37.5	10	4	US-09-620-091-42	Sequence 42, Appl
43	27	37.5	10	4	US-09-620-091-47	Sequence 47, Appl
44	26.5	36.8	10	3	US-09-186-958-12	Sequence 12, Appl
45	26.5	36.8	10	3	US-09-669-271A-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-535-852-1090
; Sequence 1090, Application US/09535852
; Patent No. 6638911
; GENERAL INFORMATION:
; APPLICANT: Blachuk, Orest W.
; APPLICANT: Symonds, James M.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND ETHODS FOR MODULATING
; FILE REFERENCE: DESMOSOMAL CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C6
; CURRENT APPLICATION NUMBER: US/09/535,852
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 2009
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1090
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cyclicized modulating agent comprising
; OTHER INFORMATION: deamoglein-2 cell adhesion recognition sequence
US-09-535-852-1090

Query Match 45.8%; Score 33; DB 4; Length 10;
Best Local Similarity 60.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGYWTIWC 10
|||
Db 1 CGYALDARGC 10

RESULT 2
US-09-535-852-1352
; Sequence 1352, Application US/09535852
; Patent No. 6638911
; GENERAL INFORMATION:
; APPLICANT: Blachuk, Orest W.
; APPLICANT: Symonds, James M.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND ETHODS FOR MODULATING
; FILE REFERENCE: DESMOSOMAL CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C6
; CURRENT APPLICATION NUMBER: US/09/535,852
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 2009
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1352
; LENGTH: 10

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cyclicized modulating agent comprising
; OTHER INFORMATION: desmocollin-1 cell adhesion recognition sequence
US-09-535-852-1352

Query Match          45.8%; Score 33; DB 4; Length 10;
Best Local Similarity 60.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGYWLTIWGC 10
    |||||
Db 1 CGYATTADGC 10

RESULT 3
US-08-520-535-12
; Sequence 12, Application US/08520535
; Patent No. 5817750
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Structural Mimics of RGD-Binding Sites
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/520,535
; FILING DATE: 28-AUG-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
US-08-520-535-12

Query Match          41.7%; Score 30; DB 2; Length 8;
Best Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGYWL 5
    ||||
Db 1 CDYWL 5

RESULT 4
US-09-079-432-12
; Sequence 12, Application US/09079432
; Patent No. 5955572
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Structural Mimics of RGD-Binding Sites
; NUMBER OF SEQUENCES: 28
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,432
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/520,535
; FILING DATE: 28-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
US-09-079-432-12

Query Match          41.7%; Score 30; DB 2; Length 8;
Best Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGYWL 5
    ||||
Db 1 CDYWL 5

RESULT 5
US-08-986-234-92
; Sequence 92, Application US/08986234
; Patent No. 5981706
; GENERAL INFORMATION:
; APPLICANT: Wallen, et al.
; TITLE OF INVENTION: Methods for Synthesizing Heat Shock Protein Complexes
; FILE REFERENCE: UNME-0008-1
; CURRENT APPLICATION NUMBER: US/08/986,234
; CURRENT FILING DATE: 1997-12-05
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 92
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-08-986-234-92

Query Match          41.7%; Score 30; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 IWGC 10
    |||||
Db 2 IWGC 5

RESULT 6
US-09-311-784A-374
; Sequence 374, Application US/09311784A
; Patent No. 6534482
```


GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epiimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-2022.01
; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR FILING DATE: 1998-05-15
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 374
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HIV1 ENV 69 (peptide 25.0113)
US-09-311-784A-374

Query Match 41.7%; Score 30; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 IWGC 10
Db 1 IWGC 4

RESULT 7
US-09-790-497A-1
; Sequence 1, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 1996-09-30
; PRIOR FILING DATE: 1993-11-22
; PRIOR FILING DATE: 1993-11-22
; PRIOR FILING DATE: 1993-03-08
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-790-497A-1

Query Match 41.7%; Score 30; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 IWGC 10
Db 1 IWGC 4

Db 1 IWGC 4

RESULT 8

US-09-108-709-22
; Sequence 22, Application US/09108709
; Patent No. 608044
; GENERAL INFORMATION:
; APPLICANT: Cotropia, Joseph P.
; TITLE OF INVENTION: Human Monoclonal Antibodies Directed Against the Transmembrane Gly
; TITLE OF INVENTION: (gp41) of Human Immunodeficiency Virus-1 (HIV-1) and Prognosis Re
; TITLE OF INVENTION: Detecting the Presence and Concentration of Antibodies Inhibiting
; TITLE OF INVENTION: Fusion-associated Epitope (GCSGKLC) in gp-41
; FILE REFERENCE: 10586/00406
; CURRENT APPLICATION NUMBER: US/09/108,709
; CURRENT FILING DATE: 1998-07-01
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: amino acids 600-609 according to the Gnann
; OTHER INFORMATION: numbering system
US-09-108-709-22

Query Match 41.7%; Score 30; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 IWGC 10
Db 1 IWGC 4

RESULT 9

US-09-790-497A-110
; Sequence 110, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 1996-09-30
; PRIOR FILING DATE: 1993-11-22
; PRIOR FILING DATE: 1993-03-08
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 110
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-790-497A-110

Query Match 41.7%; Score 30; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;

```
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 IWGC 10
   ||||
Db 1 IWGC 4

RESULT 10
US-07-946-237-4
; Sequence 4, Application US/07946237
; Patent No. 5348874
; GENERAL INFORMATION:
; APPLICANT: Savakis, Charalambos
; APPLICANT: Franz, Gerald H.
; APPLICANT: Loukeris, Athanasios
; TITLE OF INVENTION: Eukaryotic Transposable Element
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,237
; FILING DATE: 19920914
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: David E. Brook
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: BTT92-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: AMINO ACID
; MOLECULE TYPE: peptide
US-07-946-237-4

Query Match 40.3%; Score 29; DB 1; Length 5;
Best Local Similarity 75.0%; Pred. No. 3.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 IWGC 10
   ||||
Db 2 VWGC 5

RESULT 11
US-08-530-566-10
; Sequence 10, Application US/08530566
; Patent No. 5840865
; GENERAL INFORMATION:
; APPLICANT: Savakis, Charalambos
; APPLICANT: Franz, Gerald H.
; APPLICANT: Loukeris, Athanasios
; TITLE OF INVENTION: Eukaryotic Transposable Element
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 IWGC 10
   ||||
Db 2 VWGC 5

RESULT 12
US-09-195-726-10
; Sequence 10, Application US/09195726
; Patent No. 6159717
; GENERAL INFORMATION:
; APPLICANT: Savakis, Charalambos
; APPLICANT: Franz, Gerald H.
; APPLICANT: Loukeris, Athanasios
; TITLE OF INVENTION: Eukaryotic Transposable Element
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02421
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/195,726
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US 08/530,566
; APPLICATION NUMBER:
; FILING DATE: 20-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/239,765
; FILING DATE: 09-MAY-1994
; PRIOR APPLICATION DATA:
```

```
/ APPLICATION NUMBER: US 07/946,237
/ FILING DATE: 14-SEP-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Carroll, Alice O.
/ REGISTRATION NUMBER: 33,542
/ REFERENCE/DOCKET NUMBER: IMB92-01ZAZ
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781) 861-6240
/ TELEFAX: (781) 861-9540
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 5 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-09-195-726-10

Query Match 40.3%; Score 29; DB 3; Length 5;
Best Local Similarity 75.0%; Pred. No. 3.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 IWGC 10
Db 2 VMGC 5

RESULT 13
US-09-067-755-10
/ Sequence 10, Application US/09067755
/ Patent No. 6225121
/ GENERAL INFORMATION:
/ APPLICANT: Savakis, Charalambos
/ APPLICANT: Franz, Gerald H.
/ APPLICANT: Loukeris, Athanasios G.
/ TITLE OF INVENTION: Eukaryotic Transposable Element
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
/ STREET: Two Militia Drive
/ CITY: Lexington
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02421
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/067,755
/ FILING DATE: 27-APR-1998
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/530,566
/ FILING DATE: 20-SEP-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/239,765
/ FILING DATE: 09-MAY-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/946,237
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Carroll, Alice O.
/ REGISTRATION NUMBER: 33,542
/ REFERENCE/DOCKET NUMBER: IMB92-01ZAZ
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781) 861-6240
/ TELEFAX: (781) 861-9540
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 5 amino acids
```

```
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-09-067-755-10

Query Match 40.3%; Score 29; DB 3; Length 5;
Best Local Similarity 75.0%; Pred. No. 3.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 IWGC 10
Db 2 VMGC 5

RESULT 14
US-08-239-765C-4
/ Sequence 4, Application US/08239765C
/ Patent No. 6469228
/ GENERAL INFORMATION:
/ APPLICANT: Savakis, Charalambos
/ APPLICANT: Franz, Gerald H.
/ APPLICANT: Loukeris, Athanasios
/ TITLE OF INVENTION: Eukaryotic Transposable Element
/ FILE REFERENCE: 18747/1130
/ CURRENT APPLICATION NUMBER: US/08/239,765C
/ CURRENT FILING DATE: 1994-05-09
/ PRIOR APPLICATION NUMBER: 07/946,237
/ PRIOR FILING DATE: 1992-09-14
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 4
/ LENGTH: 5
/ TYPE: PRT
/ ORGANISM: Unknown
/ FEATURE:
/ OTHER INFORMATION: Conserved amino acid sequence of the TC-1 family of transposable
US-08-239-765C-4

Query Match 40.3%; Score 29; DB 4; Length 5;
Best Local Similarity 75.0%; Pred. No. 3.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 IWGC 10
Db 2 VMGC 5

RESULT 15
US-09-069-827A-94
/ Sequence 94, Application US/09069827A
/ Patent No. 6617114
/ GENERAL INFORMATION:
/ APPLICANT: FOWLKES, Dana M
/ KAY, Brian K
/ FRELINGER, Jeffrey A
/ HYDE-DERUYSCHE, Robin P
/ TITLE OF INVENTION: IDENTIFICATION OF DRUGS USING
/ COMPLEMENTARY COMBINATORIAL LIBRARIES
/ NUMBER OF SEQUENCES: 178
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BROWDY AND NEWMARK, P.L.L.C.
/ STREET: 624 Ninth Street N.W., Suite 300
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20001
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
;
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/069,827A
; FILING DATE: 30-Apr-1998
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/050,359
; FILING DATE: 31-MAR-1998
; APPLICATION NUMBER: PCT/US97/19638
; FILING DATE: 31-OCT-1997
; APPLICATION NUMBER: US 08/740,671
; FILING DATE: 31-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Iver P
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: FOWLKES-4C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 94:
US-09-069-827A-94

Query Match      40.3%; Score 29; DB 4; Length 7;
Best Local Similarity 57.1%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      3 YWLTING 9
Db      1 YWMPDWG 7

Search completed: January 3, 2005, 16:11:56
Job time : 38 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:10:30 ; Search time 139 Seconds
(without alignments)
25.880 Million cell updates/sec

Title: SEQ33

Perfect score: 72

Sequence: 1 cgywltiwcg 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 182644

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	100.0	10	13	US-10-046-922-35
2	54	75.0	10	13	US-10-046-922-34
3	50	69.4	10	13	US-10-046-922-73
4	33	45.8	7	13	US-10-046-922-67
5	33	45.8	8	13	US-10-046-922-68
6	33	45.8	10	17	US-10-654-578-1090
7	33	45.8	10	17	US-10-654-578-1352
8	32	44.4	10	8	US-08-821-739A-90
9	32	44.4	10	14	US-10-133-210-39
10	32	44.4	10	14	US-10-133-210-69
11	31	43.1	9	13	US-10-046-922-36
12	31	43.1	10	10	US-09-572-404B-1454
13	30	41.7	9	8	US-08-821-739A-78

14	41.7	9	9	US-09-832-723-104	Sequence 104, App
15	30	41.7	9	US-09-894-018-80	Sequence 80, Appl
16	30	41.7	9	US-09-894-018-185	Sequence 185, App
17	30	41.7	9	US-10-303-331-104	Sequence 104, App
18	30	41.7	9	US-10-371-525-374	Sequence 374, App
19	30	41.7	9	US-10-371-069-374	Sequence 374, App
20	30	41.7	9	US-10-371-645-374	Sequence 374, App
21	30	41.7	9	US-10-371-260-374	Sequence 374, App
22	30	41.7	9	US-10-474-960A-80	Sequence 80, Appl
23	30	41.7	9	US-10-474-960A-185	Sequence 185, App
24	30	41.7	10	US-10-462-452-707	Sequence 707, App
25	30	41.7	10	US-10-601-953-812	Sequence 812, App
26	30	41.7	10	US-10-322-266-708	Sequence 708, App
27	29	40.3	5	US-08-239-765B-4	Sequence 4, Appl
28	29	40.3	9	US-09-894-018-154	Sequence 154, App
29	29	40.3	9	US-10-133-210-8	Sequence 8, Appl
30	29	40.3	9	US-10-371-525-348	Sequence 348, App
31	29	40.3	9	US-10-371-069-348	Sequence 348, App
32	29	40.3	9	US-10-371-645-348	Sequence 348, App
33	29	40.3	9	US-10-371-260-348	Sequence 348, App
34	29	40.3	9	US-10-182-252A-180	Sequence 180, App
35	29	40.3	9	US-10-182-252A-181	Sequence 181, App
36	29	40.3	9	US-10-182-252A-305	Sequence 305, App
37	29	40.3	9	US-10-182-252A-792	Sequence 792, App
38	29	40.3	9	US-10-182-252A-793	Sequence 793, App
39	29	40.3	9	US-10-182-252A-837	Sequence 837, App
40	29	40.3	9	US-10-182-252A-838	Sequence 838, App
41	29	40.3	9	US-10-182-252A-1271	Sequence 1271, Ap
42	29	40.3	9	US-10-474-960A-154	Sequence 154, App
43	29	40.3	10	US-08-821-739A-92	Sequence 92, Appl
44	29	40.3	10	US-09-947-925A-21	Sequence 21, Appl
45	29	40.3	10	US-10-094-401-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-10-046-922-35
; Sequence 35, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Karl
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/3708A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 35
; LENGTH: 10
; TYPE: PRT
; ORGANISM: isolated peptide
US-10-046-922-35

Query Match 100.0%; Score 72; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0026; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 CGYWLTIWGC 10
Db 1 CGYWLTIWGC 10

RESULT 2
US-10-046-922-34
; Sequence 34, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Karl
; APPLICANT: Koivunen, Erkki

; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 10
; TYPE: PRT
; ORGANISM: isolated peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(1)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (10)..(10)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-34

Query Match 75.0%; Score 54; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GYWLTIWG 9
Db |||||

RESULT 3
US-10-046-922-73
; Sequence 73, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73
; LENGTH: 10
; TYPE: PRT
; ORGANISM: peptide library
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)..(7)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (9)..(9)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-73

Query Match 69.4%; Score 50; DB 13; Length 10;
Best Local Similarity 60.0%; Pred. No. 2.5;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGYWLTIWGC 10
Db |||||

RESULT 4
US-10-046-922-67
; Sequence 67, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime

; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 67
; LENGTH: 7
; TYPE: PRT
; ORGANISM: peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)..(6)
; OTHER INFORMATION: X at position 4-6 is any amino acid
US-10-046-922-67

Query Match 45.8%; Score 33; DB 13; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.5e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GYWLTIW 8
Db |||||

RESULT 5
US-10-046-922-68
; Sequence 68, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 68
; LENGTH: 8
; TYPE: PRT
; ORGANISM: peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)..(6)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (8)..(8)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-68

Query Match 45.8%; Score 33; DB 13; Length 8;
Best Local Similarity 57.1%; Pred. No. 1.5e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GYWLTIW 8
Db |||||

RESULT 6
US-10-654-578-1090
; Sequence 1090, Application US/10654578
; Publication No. US20040229811A1
; GENERAL INFORMATION:
; APPLICANT: Blachuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: DESMOSOMAL CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C10
; CURRENT APPLICATION NUMBER: US/10/654,578

```
; CURRENT FILING DATE: 2003-09-03
; NUMBER OF SEQ ID NOS: 2009
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1090
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cyclicized modulating agent comprising
; OTHER INFORMATION: desmoglein-2 cell adhesion recognition sequence
US-10-654-578-1090

Query Match      45.8%; Score 33; DB 17; Length 10;
Best Local Similarity 60.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 CGYWLTIWGC 10
Db      1 CGYALDARGC 10

RESULT 7
US-10-654-578-1352
; Sequence 1352, Application US/10654578
; Publication No. US20040229811A1
; GENERAL INFORMATION:
; APPLICANT: Blachuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: DESMOSOMAL CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C10
; CURRENT APPLICATION NUMBER: US/10/654,578
; CURRENT FILING DATE: 2003-09-03
; NUMBER OF SEQ ID NOS: 2009
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1352
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cyclicized modulating agent comprising
; OTHER INFORMATION: desmocollin-1 cell adhesion recognition sequence
US-10-654-578-1352

Query Match      45.8%; Score 33; DB 17; Length 10;
Best Local Similarity 60.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 CGYWLTIWGC 10
Db      1 CGYATTADGC 10

RESULT 8
US-08-821-739A-90
; Sequence 90, Application US/08821739A
; Publication No. US20020168374A1
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding Peptides and Their Uses
; FILE REFERENCE: 2060.005000A
; CURRENT APPLICATION NUMBER: US/08/821,739A
; CURRENT FILING DATE: 1999-03-20
; PRIOR APPLICATION NUMBER: 60/013,833
; PRIOR FILING DATE: 1996-03-21
; PRIOR APPLICATION NUMBER: 08/589,107
; PRIOR FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 08/451,913
; PRIOR FILING DATE: 1995-05-26
```

```
; PRIOR APPLICATION NUMBER: 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: 08/186,266
; PRIOR FILING DATE: 1994-01-25
; PRIOR APPLICATION NUMBER: 08/159,339
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: 08/103,396
; PRIOR FILING DATE: 1993-08-06
; PRIOR APPLICATION NUMBER: 08/027,746
; PRIOR FILING DATE: 1993-03-05
; PRIOR APPLICATION NUMBER: 07/926,666
; PRIOR FILING DATE: 1992-08-07
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 90
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-08-821-739A-90

Query Match      44.4%; Score 32; DB 8; Length 10;
Best Local Similarity 83.3%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 LTIWGC 10
Db      2 LGIWGC 7

RESULT 9
US-10-133-210-39
; Sequence 39, Application US/10133210
; Publication No. US20030103964A1
; GENERAL INFORMATION:
; APPLICANT: Delisi, Charles
; APPLICANT: Berzofsky, Jay
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Vaccaro, Dennis
; APPLICANT: Weng, Zhiping
; APPLICANT: Zhang, Chao
; TITLE OF INVENTION: METHODS FOR DESIGNING MOLECULAR CONJUGATES AND
; TITLE OF INVENTION: COMPOSITIONS THEREOF
; FILE REFERENCE: BU-035AX
; CURRENT APPLICATION NUMBER: US/10/133,210
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 281
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-133-210-39

Query Match      44.4%; Score 32; DB 14; Length 10;
Best Local Similarity 83.3%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 LTIWGC 10
Db      2 LGIWGC 7

RESULT 10
US-10-133-210-69
; Sequence 69, Application US/10133210
; Publication No. US20030103964A1
; GENERAL INFORMATION:
; APPLICANT: Delisi, Charles
; APPLICANT: Berzofsky, Jay
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Vaccaro, Dennis
```

```
; APPLICANT: Weng, Zhiping
; APPLICANT: Zhang, Chao
; TITLE OF INVENTION: METHODS FOR DESIGNING MOLECULAR CONJUGATES AND
; FILE REFERENCE: BU-035AX
; CURRENT APPLICATION NUMBER: US/10/133,210
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 281
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 69
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-133-210-69

Query Match      44.4%; Score 32; DB 14; Length 10;
Best Local Similarity 83.3%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 LTIWGC 10
        | | | | |
Db      2 LGIWGC 7

RESULT 11
US-10-046-922-36
; Sequence 36, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 36
; LENGTH: 9
; TYPE: PRT
; ORGANISM: peptide
US-10-046-922-36

Query Match      43.1%; Score 31; DB 13; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.5e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 GYWLTIW 8
        | | | | |
Db      2 GYWDWTW 8

RESULT 12
US-09-572-404B-1454
; Sequence 1454, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: Protpatent version 1.0
; SEQ ID NO 1454
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in P2RY5 at 139-148 and may interact with Sequen
```

```
; OTHER INFORMATION: in this patent.
US-09-572-404B-1454

Query Match      43.1%; Score 31; DB 10; Length 10;
Best Local Similarity 62.5%; Pred. No. 9.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GYWLTIW 9
        | | | | |
Db      1 GWLTIWIG 8

RESULT 13
US-08-821-739A-78
; Sequence 78, Application US/08821739A
; Publication No. US20020168374A1
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding Peptides and Their Uses
; FILE REFERENCE: 2060.005000A
; CURRENT APPLICATION NUMBER: US/08/821,739A
; CURRENT FILING DATE: 1999-03-20
; PRIOR APPLICATION NUMBER: 60/013,833
; PRIOR FILING DATE: 1996-03-21
; PRIOR APPLICATION NUMBER: 08/589,107
; PRIOR FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 08/451,913
; PRIOR FILING DATE: 1995-05-26
; PRIOR APPLICATION NUMBER: 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: 08/186,266
; PRIOR FILING DATE: 1994-01-25
; PRIOR APPLICATION NUMBER: 08/159,339
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: 08/103,396
; PRIOR FILING DATE: 1993-08-06
; PRIOR APPLICATION NUMBER: 08/027,746
; PRIOR FILING DATE: 1993-03-05
; PRIOR APPLICATION NUMBER: 07/926,666
; PRIOR FILING DATE: 1992-08-07
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 78
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-821-739A-78

Query Match      41.7%; Score 30; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 IWGC 10
        | | | | |
Db      1 IWGC 4

RESULT 14
US-09-832-723-104
; Sequence 104, Application US/09832723
; Patent No. US20020098524A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; APPLICANT: Chen, Yiyou
; APPLICANT: Murray, Christopher J.
; APPLICANT: Tijerina, Pilar
; TITLE OF INVENTION: METHODS FOR SELECTIVE TARGETING
; FILE REFERENCE: GC617-2
; CURRENT APPLICATION NUMBER: US/09/832,723
; CURRENT FILING DATE: 2001-04-11
```


; PRIOR APPLICATION NUMBER: US 60/197,259
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptides screened from a phage display random
; OTHER INFORMATION: peptide library
US-09-832-723-104

Query Match 41.7%; Score 30; DB 9; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.5e+06;
Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGYWLTIW 8
| | | | |
Db 1 CTLWPTFW 8

RESULT 15
US-09-894-018-80
; Sequence 80, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligopeptide
US-09-894-018-80

Query Match 41.7%; Score 30; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 IWGC 10
| | | | |
Db 1 IWGC 4

Search completed: January 3, 2005, 16:22:31
Job time : 140 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 15:54:23 ; Search time 43 Seconds
(without alignments)
22.376 Million cell updates/sec

Title: SEQ33

Perfect score: 72

Sequence: 1 cgywltiwc 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0

Maximum-DB-seq-length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	33.3	10	2 S71868	glutathione transf
2	24	33.3	10	2 A35556	hypothetical prote
3	22	30.6	7	2 PT0586	T-cell receptor be
4	22	30.6	10	2 T17075	cytochrome-c oxida
5	21	29.2	10	2 E41946	T-cell receptor ga
6	20	27.8	5	2 JH0253	gut pentapeptide -
7	20	27.8	6	2 F41946	T-cell receptor ga
8	20	27.8	8	2 S19288	acylase - Kluuvera
9	19	26.4	9	2 PT0324	Ig heavy chain CRD
10	19	26.4	10	2 PT0289	Ig heavy chain CRD
11	19	26.4	10	2 E49033	T-cell receptor ga
12	19	26.4	10	2 F49033	T-cell receptor ga
13	19	26.4	10	2 C41946	T-cell receptor ga
14	18	25.0	8	2 JS0315	leucokinin V - Mad
15	18	25.0	10	2 PT0230	Ig heavy chain CDR
16	18	25.0	10	2 PH0923	T-cell receptor be
17	18	25.0	10	2 F33932	Ig mu chain J regi
18	18	25.0	10	2 S14943	UGA3 leader peptid
19	17	23.6	4	2 B53284	T-cell receptor be
20	17	23.6	6	2 PT0829	T-cell receptor be
21	17	23.6	6	2 PT0637	T-cell receptor be
22	17	23.6	6	2 A61068	locustakinin - mig
23	17	23.6	6	4 I79564	hypothetical TGL3
24	17	23.6	7	2 PT0628	T-cell receptor be
25	17	23.6	7	2 PT0642	T-cell receptor be
26	17	23.6	7	2 PT0722	T-cell receptor be
27	17	23.6	7	2 PT0728	T-cell receptor be
28	17	23.6	7	2 PX0008	glucuronosyltransf
29	17	23.6	7	2 B48394	major fat-globule

30 17 23.6 7 2 PD0029
31 17 23.6 7 2 S57274
32 17 23.6 7 2 S33244
33 17 23.6 7 2 S33245
34 17 23.6 7 2 S33246
35 17 23.6 8 2 PT0724
36 17 23.6 8 2 JS0316
37 17 23.6 8 2 JS0317
38 17 23.6 8 2 JS0318
39 17 23.6 8 2 A31570
40 17 23.6 9 2 A24244
41 17 23.6 9 2 PT0634
42 17 23.6 9 2 PT0562
43 17 23.6 9 2 A60522
44 17 23.6 10 2 B33995
45 17 23.6 10 2 S08997

ALIGNMENTS

RESULT 1

S71868

glutathione transferase (EC 2.5.1.18) class mu 4 - pig (fragment)

N;Alternate names: glutathione S-transferase class mu 4

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 19-Mar-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004

C;Accession: S71868

R;Rouimi, P.; Anglade, P.; Debrauwer, L.; Tulliez, J.

Biochem. J. 317, 879-884, 1996

A;Title: Characterization of pig liver glutathione S-transferases using HPLC-electrospra

A;Reference number: S71864; MUID:96332484; PMID:8760377

A;Accession: S71868

A;Molecule type: protein

A;Residues: 1-10 <ROU>

C;Cross-references: UNIPROT:Q7M3E8

C;Comment: At least five species-independent classes of cytosolic glutathione transferases

s mitochondrial form are known.

C;Complex: dimer

C;Function:

A;Description: catalyzes the nucleophilic conjugation of intracellular glutathione to a

A;Pathway: detoxification; xenobiotics metabolism

A;Note: increased hydrophilicity of GSH-conjugates facilitates their further metabolism

es of damage

C;Superfamily: glutathione transferase

C;Keywords: dimer; transferase

Query Match 33.3%; Score 24; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 9.9e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GYW 4

DB 4 GYW 6

RESULT 2

A35556

hypothetical protein (ODC region) - human

C;Species: Homo sapiens (man)

C;Date: 21-Sep-1990 #sequence_revision 21-Sep-1990 #text_change 30-Sep-1993

C;Accession: A35556

R;Moehner, J.A.; Gilbert, J.D.; Skunca, M.; Doseescu, J.; Almodovar, K.M.; Luk, G.D.

J. Biol. Chem. 265, 4884-4892, 1990

A;Title: Isolation and expression of a human ornithine decarboxylase gene.

A;Reference number: A35556; MUID:90202959; PMID:2318872

A;Accession: A35556

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-10 <MOS>

A;Cross-references: GB:J05271

Query Match 33.3%; Score 24; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 9.9e+02;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGYW 4
Db 5 CGAW 8

RESULT 3
PT0586
T-cell receptor beta chain V-D-J region (141-1CN) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0586; PT0592
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0586
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <PEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c (clones 141-1CN and 141-1CD)
C;Keywords: T-cell receptor

Query Match 30.6%; Score 22; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIWG 9
Db 3 SIWG 6

RESULT 4
T17075
cytochrome-c oxidase (EC 1.9.3.1) chain I - Chamaeleo fischeri mitochondrion (fragment)
C;Species: mitochondrion Chamaeleo fischeri
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17075
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Papenfuss, T.J.
J. Mol. Evol. 44, 660-674, 1997
A;Title: Evolutionary shifts in three major structural features of the mitochondrial gene
A;Reference number: Z18674; MUID:97315309; PMID:9169559
A;Accession: T17075
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-10 <MAC>
A;Cross-references: UNIPROT:O79912; EMBL:U82688; NID:g3603112; PID:g3603115; PIDN:AAC622
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

Query Match 30.6%; Score 22; DB 2; Length 10;
Best Local Similarity 60.0%; Pred. No. 2e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 WLTW 8
Db 2 WLLRW 6

RESULT 5
E41946
T-cell receptor gamma chain (1a.9) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C;Accession: E41946
R;Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
A;Reference number: A41946; MUID:92049316; PMID:1658619

Best Local Similarity 75.0%; Pred. No. 9.9e+02;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGYW 4
Db 5 CGAW 8

RESULT 3
PT0586
T-cell receptor beta chain V-D-J region (141-1CN) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0586; PT0592
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0586
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <PEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c (clones 141-1CN and 141-1CD)
C;Keywords: T-cell receptor

Query Match 30.6%; Score 22; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIWG 9
Db 3 SIWG 6

RESULT 4
T17075
cytochrome-c oxidase (EC 1.9.3.1) chain I - Chamaeleo fischeri mitochondrion (fragment)
C;Species: mitochondrion Chamaeleo fischeri
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17075
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Papenfuss, T.J.
J. Mol. Evol. 44, 660-674, 1997
A;Title: Evolutionary shifts in three major structural features of the mitochondrial gen
A;Reference number: Z18674; MUID:97315309; PMID:9169559
A;Accession: T17075
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-10 <MAC>
A;Cross-references: UNIPROT:O79912; EMBL:U82688; NID:g3603112; PID:g3603115; PIDN:AAC622
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

Query Match 30.6%; Score 22; DB 2; Length 10;
Best Local Similarity 60.0%; Pred. No. 2e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 WLTW 8
Db 2 WLLRW 6

RESULT 5
E41946
T-cell receptor gamma chain (1a.9) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C;Accession: E41946
R;Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
A;Reference number: A41946; MUID:92049316; PMID:1658619

A;Accession: E41946
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-10 <WHE>
C;Keywords: T-cell receptor

Query Match 29.2%; Score 21; DB 2; Length 10;
Best Local Similarity 40.0%; Pred. No. 2.8e+03;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYW 5
Db 2 CAVWI 6

RESULT 6
JH0253
gut pentapeptide - Japanese eel
C;Species: Anguilla japonica (Japanese eel)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Apr-1995
C;Accession: JH0253
R;Oesaka, T.; Ikeda, T.; Kubota, I.; Muneoka, Y.; Ando, M.
Biochem. Biophys. Res. Commun. 180, 828-832, 1991
A;Title: Structure and function of a pentapeptide isolated from the gut of the eel.
A;Reference number: JH0253; MUID:92062113; PMID:1953755
A;Accession: JH0253
A;Molecule type: protein
A;Residues: 1-5 <UES>
A;Experimental source: gut
C;Comment: This peptide increased basal tone of the circular muscle of the esophagogastr
, and of the circular muscle of the gastro-intestinal junction.

Query Match 27.8%; Score 20; DB 2; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GYW 4
Db 1 GFW 3

RESULT 7
F41946
T-cell receptor gamma chain (1a.27) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C;Accession: F41946
R;Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
A;Reference number: A41946; MUID:92049316; PMID:1658619
A;Accession: F41946
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-6 <WHE>
C;Keywords: T-cell receptor

Query Match 27.8%; Score 20; DB 2; Length 6;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYW 5
Db 2 CAVWV 6

RESULT 8
S19288
acylase - Kluyvera cryocrescens
C;Species: Kluyvera cryocrescens
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S19288
R;Martin, J.; Slade, A.; Aitken, A.; Arche, R.; Virden, R.

Biochem. J. 280, 659-662, 1991
A:Title: Chemical modification of serine at the active site of penicillin acylase from K
A:Reference number: S19288; MUID:92109664; PMID:1764029
A:Accession: S19288
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <MAR>
A:Cross-references: UNIPROT:Q7M124

Query Match 27.8%; Score 20; DB 2; Length 8;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYWL 5
| :
| :
Db 1 CNMNV 5

RESULT 9
PT0324
Ig heavy chain CRD3 region (clone J2-106C) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0324
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J
A:Reference number: PT0222; MUID:91108337; PMID:1899102
A:Accession: PT0324
A:Molecule type: DNA
A:Experimental source: B lymphocyte
A:Keywords: heterotetramer; immunoglobulin

Query Match 26.4%; Score 19; DB 2; Length 9;
Best Local Similarity 42.9%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GYWLTIW 8
| :
| :
Db 3 GYGESYW 9

RESULT 10
PT0289
Ig heavy chain CRD3 region (clone 4-109) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0289
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J
A:Reference number: PT0222; MUID:91108337; PMID:1899102
A:Accession: PT0289
A:Molecule type: DNA
A:Residues: 1-10 <YAM>
A:Experimental source: B lymphocyte
A:Keywords: heterotetramer; immunoglobulin

Query Match 26.4%; Score 19; DB 2; Length 10;
Best Local Similarity 33.3%; Pred. No. 5.4e+03;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 WLTIWG 9
| :
| :
Db 4 WISMG 9

RESULT 11
E49033
T-cell receptor gamma chain V-D-J region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 19-Dec-1993 #sequence_revision 25-Aug-1995 #text_change 21-Jul-2000

C:Accession: E49033; D49033
R:Morita, C.T.; Verma, S.; Aparicio, P.; Martinez, C.; Spits, H.; Brenner, M.B.
Eur. J. Immunol. 21, 2999-3007, 1991
A:Title: Functionally distinct subsets of human gamma/delta T cells.
A:Reference number: A49033; MUID:92083926; PMID:1684157
A:Accession: E49033
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-10 <MOR>
A:Cross-references: GB:S72587; NID:9240696; PIDN:AAB20630.1; PID:9240697
A>Note: sequence extracted from NCBI backbone (NCBIN:72591, NCBIIP:72595)
A:Accession: D49033
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-10 <MOR2>
A:Cross-references: GB:S72587; NID:9240696; PIDN:AAB20630.1; PID:9240697
A>Note: sequence extracted from NCBI backbone (NCBIN:72587, NCBIIP:72589)
C:Keywords: T-cell receptor

Query Match 26.4%; Score 19; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 5.4e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYW 4
| :
| :
Db 1 CALW 4

RESULT 12
F49033
T-cell receptor gamma chain V-D-J region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 19-Dec-1993 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: F49033
R:Morita, C.T.; Verma, S.; Aparicio, P.; Martinez, C.; Spits, H.; Brenner, M.B.
Eur. J. Immunol. 21, 2999-3007, 1991
A:Title: Functionally distinct subsets of human gamma/delta T cells.
A:Reference number: A49033; MUID:92083926; PMID:1684157
A:Accession: F49033
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-10 <MOR>
A:Cross-references: GB:S72605; NID:9240700; PIDN:AAB20632.1; PID:9240701
A>Note: sequence extracted from NCBI backbone (NCBIN:72605, NCBIIP:72606)
C:Keywords: T-cell receptor

Query Match 26.4%; Score 19; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 5.4e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYW 4
| :
| :
Db 1 CALW 4

RESULT 13
C41946
T-cell receptor gamma chain (1t.50) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C:Accession: C41946
R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
A:Reference number: A41946; MUID:92049316; PMID:1658619
A:Accession: C41946
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-10 <WHE>
C:Keywords: T-cell receptor

Query Match 26.4%; Score 19; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 5.4e+03;

Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYW 4
| |
Db 2 CAVW 5

RESULT 14

JS0315
leucokinin V - Madeira cockroach
C;Species: Leucophaea maderae (Madeira cockroach)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: JS0315
R;Holman, G.M.; Cook, B.J.; Nachman, R.J.
Comp. Biochem. Physiol. C 88, 27-30, 1987
A;Title: Isolation, primary structure, and synthesis of leucokinins V and VI: myotropic
A;Reference number: JS0315
A;Accession: JS0315
A;Molecule type: protein
A;Residues: 1-8 <HOL>
A;Cross-references: UNIPROT:P19987
C;Comment: Leucokinins, a family of cephalomyotropic peptides, stimulate contractile act
C;Keywords: amidated carboxyl end; cephalomyotropic peptide
P;8/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 25.0%; Score 18; DB 2; Length 8;
Best Local Similarity 37.5%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 GYWLITWG 9
| : : | |
Db 1 GSGFSSWG 8

RESULT 15

PT0230
Ig heavy chain CDR3 region (clone 1-118A) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0230
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0230
A;Molecule type: DNA
A;Residues: 1-10 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 25.0%; Score 18; DB 2; Length 10;
Best Local Similarity 60.0%; Pred. No. 7.5e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LITWG 9
: : : | |
Db 3 ITITG 7

Search completed: January 3, 2005, 16:11:14
Job time : 45 secs


```

RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22169767; PubMed=12182400;
RA  Townsend T., Larson A.;
RT  "Molecular phylogenetics and mitochondrial genomic evolution in the
RL  chamaeleonidae (Reptilia, Squamata).";
RL  Mol. Phylogenet. Evol. 23:22-36(2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Townsend T.M., Larson A.L.;
RL  Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AF448755; AAL90547.1; -.
DR  GO: GO:0005739; C:mitochondrion; IEA.
KW  Mitochondrion.
FT  NON TER 10 10
SQ  SEQUENCE 10 AA; 1276 MW; 5E218E2733772727 CRC64;

Query Match 37.5%; Score 27; DB 2; Length 10;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WLTW 8
Db 2 WLLW 6

RESULT 3
Q70Y57
ID Q70Y57 PRELIMINARY; PRT; 8 AA.
AC Q70Y57;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ribosomal protein (Fragment).
GN Name=rps16;
OS Fuerstia africana.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiaceae; Nepetoideae; Ocimeae; Fuerstia.
OX NCBI_TaxID=204226;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15019625;
RA Paton A., Springate D.A., Sudde S., Otieno D., Grayer R., Harley M.M.,
RA Willis F., Simonds M.S.J., Powell M.P., Savolainen V.;
RT "Phylogeny and evolution of basilis and allies (Ocimeae, Labiatae)
RT based on three plastid DNA regions.";
RL Mol. Phylogenet. Evol. 31:277-299(2004).
DR EMBL: AJ505427; CAD45547.1; -.
DR GO: GO:0003735; F:Structural constituent of ribosome; IEA.
KW Ribosomal protein.
FT NON TER 1 1
FT NON TER 8 8
SQ SEQUENCE 8 AA; 916 MW; DABEAB58637041B5 CRC64;

Query Match 36.1%; Score 26; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 TING 9
Db 2 TING 5

RESULT 4
CAD45547
ID CAD45547 PRELIMINARY; PRT; 8 AA.
AC CAD45547;
DT 14-MAR-2004 (TrEMBLrel. 27, Created)
DT 14-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 14-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Ribosomal protein (Fragment).

```

```

GN RPS16.
OS Fuerstia africana.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiaceae; Nepetoideae; Ocimeae; Fuerstia.
OX NCBI_TaxID=204226;
RN [1]
RP SEQUENCE FROM N.A.
RA Paton A., Springate D.A., Sudde S., Otieno D., Grayer R., Harley M.M.,
RA Willis F., Simonds M.S.J., Powell M.P., Savolainen V.;
RT "Phylogeny and evolution of basilis and allies (Ocimeae, Labiatae)
RT based on three plastid DNA regions.";
RL Mol. Phylogenet. Evol. 31:277-299(2004).
DR EMBL: AJ505427; CAD45547.1; -.
DR GO: GO:0003735; F:Structural constituent of ribosome; IEA.
KW Ribosomal protein.
FT NON TER 1 1
FT NON TER 8 8
SQ SEQUENCE 8 AA; 916 MW; DABEAB58637041B5 CRC64;

Query Match 36.1%; Score 26; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 TING 9
Db 2 TING 5

RESULT 5
Q7M3E8
ID Q7M3E8 PRELIMINARY; PRT; 10 AA.
AC Q7M3E8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glutathione transferase (EC 2.5.1.18) class mu 4 (Fragment).
OS Sus scrofa domestica (domestic pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9825;
RN [1]
RP SEQUENCE.
RA Rouimi P., Anglade P., Debrauwer L., Tulliez J.;
RT "Characterization of pig liver glutathione S-transferases using HPLC-
RT electrospray-ionization mass spectrometry.";
RL Biochem. J. 317:879-884(1996).
DR PIR: S71868; S71868.
DR GO: GO:0004364; F:glutathione transferase activity; IEA.
FT NON TER 10 10
SQ SEQUENCE 10 AA; 1223 MW; 5E16395AB36B5877 CRC64;

Query Match 33.3%; Score 24; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.8e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GYW 4
Db 4 GYW 6

RESULT 6
Q8SHN1
ID Q8SHN1 PRELIMINARY; PRT; 10 AA.
AC Q8SHN1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN Name=COI;
OS Bradypodion tavetanum (Dwarf Fischer's chameleon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```



```

Q85DB8
ID Q85DB8 PRELIMINARY; PRT; 9 AA.
AC Q85DB8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome oxidase subunit III (Fragment).
GN Name=COIII;
OS Lepilemur edwardsi (Milne-Edwards's sportive lemur).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Megaladapidae; Lepilemur.
OX NCBI_TaxID=122230;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22631663; PubMed=12719521;
RA Pastorini J., Thalmann U., Martin R.D.;
RT "A molecular approach to comparative phylogeography of extant Malagasy
lemons.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5879-5884 (2003).
DR EMBL; AF224595; AAF33644.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 32.6%; Score 23.5; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 3 YWLTWG 9
Db ||||
5 YW---WG 8

RESULT 11
Q94NA9
ID Q94NA9 PRELIMINARY; PRT; 9 AA.
AC Q94NA9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit III (Fragment).
GN Name=COIII;
OS Daubentonia madagascariensis (Aye-aye).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Daubentoniidae;
OC Daubentonia.
OX NCBI_TaxID=31869;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21184272; PubMed=11286490;
RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
RT "Molecular phylogeny of the lemur family cheirogaleidae (primates)
based on mitochondrial DNA sequences.";
RL Mol. Phylogenet. Evol. 19:45-56 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22631663; PubMed=12719521;
RA Pastorini J., Thalmann U., Martin R.D.;
RT "A molecular approach to comparative phylogeography of extant Malagasy
lemons.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5879-5884 (2003).
DR EMBL; AF224636; AAK70599.1; -.
DR EMBL; AF224637; AAK70599.1; -.
DR EMBL; AF224638; AAK70603.1; -.
DR EMBL; AF224639; AAK70607.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 32.6%; Score 23.5; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 3 YWLTWG 9
Db ||||
5 YW---WG 8

RESULT 12
Q94NB0
ID Q94NB0 PRELIMINARY; PRT; 9 AA.
AC Q94NB0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit III (fragment).
GN Name=COIII;
OS Microcebus rufus (Brown mouse lemur).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
OC Microcebus.
OX NCBI_TaxID=122232;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21184272; PubMed=11286490;
RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
RT "Molecular phylogeny of the lemur family cheirogaleidae (primates)
based on mitochondrial DNA sequences.";
RL Mol. Phylogenet. Evol. 19:45-56 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22631663; PubMed=12719521;
RA Pastorini J., Thalmann U., Martin R.D.;
RT "A molecular approach to comparative phylogeography of extant Malagasy
lemons.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5879-5884 (2003).
DR EMBL; AF224636; AAK70599.1; -.
DR EMBL; AF224637; AAK70599.1; -.
DR EMBL; AF224638; AAK70603.1; -.
DR EMBL; AF224639; AAK70607.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 32.6%; Score 23.5; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 3 YWLTWG 9
Db ||||
5 YW---WG 8

RESULT 13
Q94NB1
ID Q94NB1 PRELIMINARY; PRT; 9 AA.
AC Q94NB1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit III (fragment).
GN Name=COIII;
OS Microcebus ravelobensis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

DR GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 32.6%; Score 23.5; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 3 YWLTWG 9
Db ||||
5 YW---WG 8

RESULT 12
Q94NB0
ID Q94NB0 PRELIMINARY; PRT; 9 AA.
AC Q94NB0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit III (fragment).
GN Name=COIII;
OS Microcebus rufus (Brown mouse lemur).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
OC Microcebus.
OX NCBI_TaxID=122232;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21184272; PubMed=11286490;
RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
RT "Molecular phylogeny of the lemur family cheirogaleidae (primates)
based on mitochondrial DNA sequences.";
RL Mol. Phylogenet. Evol. 19:45-56 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22631663; PubMed=12719521;
RA Pastorini J., Thalmann U., Martin R.D.;
RT "A molecular approach to comparative phylogeography of extant Malagasy
lemons.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5879-5884 (2003).
DR EMBL; AF224636; AAK70599.1; -.
DR EMBL; AF224637; AAK70599.1; -.
DR EMBL; AF224638; AAK70603.1; -.
DR EMBL; AF224639; AAK70607.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 32.6%; Score 23.5; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 3 YWLTWG 9
Db ||||
5 YW---WG 8

RESULT 13
Q94NB1
ID Q94NB1 PRELIMINARY; PRT; 9 AA.
AC Q94NB1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit III (fragment).
GN Name=COIII;
OS Microcebus ravelobensis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
 OC Microcebus.
 OX NCBI_TaxID=122231;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21184272; PubMed=11286490;
 RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
 RT "Molecular phylogeny of the lemur family cheirogaleidae (primates)
 based on mitochondrial DNA sequences.";
 RL Mol. Phylogenet. Evol. 19:45-56(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22631663; PubMed=12719521;
 RA Pastorini J., Thalmann U., Martin R.D.;
 RT "A molecular approach to comparative phylogeography of extant Malagasy
 lemurs.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:5879-5884(2003).
 DR EMBL; AF224630; AAK70571.1; -;
 DR EMBL; AF224631; AAK70575.1; -;
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON_TER 1
 SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 32.6%; Score 23.5; DB 2; Length 9;
 Best Local Similarity 57.1%; Pred. No. 1.8e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 3 YWLTWIG 9
 DB 5 YW---WG 8

RESULT 14

O94NB2 PRELIMINARY; PRT; 9 AA.
 AC O94NB2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Cytochrome oxidase subunit III (Fragment).
 GN Names=COIII;
 OS Microcebus murinus (Lesser mouse lemur).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
 OC Microcebus.
 OX NCBI_TaxID=30608;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21184272; PubMed=11286490;
 RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
 RT "Molecular phylogeny of the lemur family cheirogaleidae (primates)
 based on mitochondrial DNA sequences.";
 RL Mol. Phylogenet. Evol. 19:45-56(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22631663; PubMed=12719521;
 RA Pastorini J., Thalmann U., Martin R.D.;
 RT "A molecular approach to comparative phylogeography of extant Malagasy
 lemurs.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:5879-5884(2003).
 DR EMBL; AF224624; AAK70547.1; -;
 DR EMBL; AF224625; AAK70551.1; -;
 DR EMBL; AF224626; AAK70555.1; -;
 DR EMBL; AF224627; AAK70559.1; -;
 DR EMBL; AF224628; AAK70563.1; -;
 DR EMBL; AF224629; AAK70567.1; -;
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON_TER 1
 SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 32.6%; Score 23.5; DB 2; Length 9;
 Best Local Similarity 57.1%; Pred. No. 1.8e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 3 YWLTWIG 9
 DB 5 YW---WG 8

Query Match 32.6%; Score 23.5; DB 2; Length 9;
 Best Local Similarity 57.1%; Pred. No. 1.8e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 3 YWLTWIG 9
 DB 5 YW---WG 8

RESULT 15

O94XE6 PRELIMINARY; PRT; 9 AA.
 AC O94XE6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cytochrome c oxidase subunit III (Fragment).
 GN Names=cox3;
 OS Tectocoris diophthalmus (cotton harlequin bug).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
 OC Panheteroptera; Pentatomomorpha; Pentatomoidea; Pentatomidae;
 OC Tectocoris.
 OX NCBI_TaxID=159956;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21396409; PubMed=11504862;
 RA Shao R., Campbell N.J., Schmidt E.R., Barker S.C.;
 RT "Increased rate of gene rearrangement in the mitochondrial genomes of
 three orders of hemipteroid insects.";
 RL Mol. Biol. Evol. 18:1828-1832(2001).
 DR EMBL; AF335990; AAK5283.1; -;
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON_TER 1
 SQ SEQUENCE 9 AA; 1206 MW; A2C563636B5041A6 CRC64;

Query Match 32.6%; Score 23.5; DB 2; Length 9;
 Best Local Similarity 57.1%; Pred. No. 1.8e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 3 YWLTWIG 9
 DB 5 YW---WG 8

Search completed: January 3, 2005, 16:10:25
 Job time : 186 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:11:20 ; Search time 113.6 Seconds
(without alignments)
31.578 Million cell updates/sec

Title: US-10-046-922-34

Perfect score: 56

Sequence: 1 XGYWLTWGX 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	10	5	Abp53931 VEGFR-3 b
2	56	100.0	10	5	Abp53932 VEGFR-3 b
3	44	78.6	13	6	Aao26093 FC region
4	44	78.6	13	8	Adj50760 Human ser
5	44	78.6	120	4	Aab62747 Human HIV
6	44	78.6	474	6	Abu30004 Protein e
7	44	78.6	492	7	Adc97318 E. faeciu
8	44	78.6	1140	4	Aae09365 Human ATP
9	44	78.6	1498	4	Aae09362 Mouse ATP
10	44	78.6	1503	2	Aay43544 A human M
11	44	78.6	1503	4	Aae09361 Human ATP
12	44	78.6	1503	4	Aae09370 Human ATP
13	44	78.6	1503	4	Aae09367 Human ATP
14	44	78.6	1503	4	Aae09368 Human ATP
15	44	78.6	1503	4	Aae09369 Human ATP
16	44	78.6	1503	4	Aae09364 Human ATP
17	44	78.6	1503	4	Aae09371 Human ATP
18	44	78.6	1503	4	Aae09363 Human ATP
19	44	78.6	1503	5	Abp52113 Homo sapi
20	43	76.8	34	4	Abb38460 Peptide #
21	43	76.8	34	4	Aam31901 Peptide #
22	43	76.8	34	4	Aam71604 Human bon
23	43	76.8	34	4	Aam59071 Human bra
24	43	76.8	34	4	Abg53286 Human liv
25	43	76.8	34	5	Abg41416 Human pep

26	43	76.8	222	2	AAY00213
27	43	76.8	222	5	ABP43432
28	43	76.8	222	6	ABU88460
29	43	76.8	222	6	ABU13711
30	43	76.8	229	7	ADH85745
31	43	76.8	266	2	AAY00212
32	43	76.8	266	5	ABP43431
33	43	76.8	266	6	ABU88459
34	43	76.8	266	6	ABU13710
35	42	75.0	14	6	AAO26134
36	42	75.0	14	8	ADJ50801
37	42	75.0	116	2	ARI15437
38	42	75.0	116	6	ABO27261
39	42	75.0	116	6	ABO27269
40	42	75.0	116	6	ABO27263
41	42	75.0	116	6	ABO27259
42	42	75.0	116	6	ABO27255
43	42	75.0	116	6	ABO27277
44	42	75.0	116	6	ABO27273
45	42	75.0	116	6	ABO27257

ALIGNMENTS

RESULT 1

ABP53931
ID ABP53931 standard; peptide; 10 AA.

XX AC ABP53931;

XX DT 09-JAN-2003 (first entry)

XX DE VEGFR-3 binding peptide SEQ ID NO:34.

XX KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
XX KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
XX KW cytotactic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
XX KW urolary; cell surface receptor; cancer; neovascularisation;
XX KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
XX KW diabetes; PDGF; platelet derived growth factor.

XX OS Homo sapiens.

OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 1 /note= "any amino acid"

FT Misc-difference 10 /note= "any amino acid"

XX WO200257299-A2.

XX 25-JUL-2002.

XX 16-JAN-2002; 2002WO-IB000099.

XX 17-JAN-2001; 2001US-0262476P.

XX (LUDW-) LUDWIG INST CANCER RES.

XX (LICN) LICENTIA LTD.

XX Alitalo K, Koivunen E, Kubo H;

XX WPI; 2002-691521/74.

XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
PT such as cancer and diseases of neovascularization.

XX Claim 12; Page 80; 149pp; English.

XX The present invention describes an isolated peptide (I) that binds to and

CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnerary activities, and can be used in gene therapy.
 CC compositions and vulnerary activities from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
 CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 XX
 SQ Sequence 10 AA;
 Query Match 100.0%; Score 56; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.077;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XGYWLTWGX 10
 Db |||||
 1 XGYWLTWGX 10
 RESULT 2
 ABP53932
 ID ABP53932 standard; peptide; 10 AA.
 XX AC ABP53932;
 XX 09-JAN-2003 (first entry)
 XX VEGFR-3 binding peptide SEQ ID NO:35.
 DE
 XX Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
 KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
 KW cytotatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
 KW vulnerary; cell surface receptor; cancer; neovascularisation;
 KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
 KW diabetes; PDGF; platelet derived growth factor.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX WO200257299-A2.
 XX 25-JUL-2002.
 XX 16-JAN-2002; 2002WO-IB000099.
 XX 17-JAN-2001; 2001US-0262476P.
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX (LICN) LICENTIA LTD.
 XX Alitalo K, Koivunen E, Kubo H;
 XX WPI; 2002-691521/74.
 XX
 XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
 PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
 PT such as cancer and diseases of neovascularization.
 XX
 XX Claim 13; Page 80; 149pp; English.
 XX
 CC The present invention describes an isolated peptide (I) that binds to and
 CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnerary activities, and can be used in gene therapy.
 CC compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 XX
 SQ Sequence 10 AA;
 Query Match 100.0%; Score 56; DB 5; Length 10;
 Best Local Similarity 80.0%; Pred. No. 0.077;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XGYWLTWGX 10
 Db :|||||
 1 CGYWLTIWGC 10
 RESULT 3
 AAO26093
 ID AAO26093 standard; peptide; 13 AA.
 XX AC AAO26093;
 XX 03-APR-2003 (first entry)
 XX Fc region binding peptide SEQ ID No 73.
 DE
 XX Immunoglobulin Fc region; binding; whole blood; plasma; transgenic milk;
 KW antibody response; half-life; stability; circulatory system.
 KW Unidentified.
 OS
 XX WO200286070-A2.
 XX 31-OCT-2002.
 XX 18-APR-2002; 2002WO-US012492.
 XX 18-APR-2001; 2001US-0284534P.
 XX (DYAX-) DYAX CORP.
 XX Rondon LJ, Wu Q, Ley AC, Stochl M, Ransohoff TC, Potter MD;
 XX WPI; 2003-201220/19.
 XX
 XX New polypeptides, useful as binding molecules for detecting, isolating or
 PT purifying immunoglobulin Fc-region polypeptides present in a solution, or
 PT for regulating or preventing an antibody response.
 XX
 XX Claim 3; Page 76; 152pp; English.
 XX
 CC The invention relates to novel isolated polypeptides comprising a
 CC sequence that binds an immunoglobulin Fc region. The polypeptides are
 CC useful as binding molecules for detecting, isolating or purifying
 CC immunoglobulin Fc-region polypeptides present in a solution, e.g. whole
 CC blood, plasma or transgenic milk. The Fc-region binding polypeptides are
 CC also useful for regulating or preventing an antibody response, or for
 CC increasing the half-life and over all stability of a therapeutic or
 CC diagnostic compound that is administered to or enters the circulatory
 CC system of an individual. This sequence represents an Fc region binding
 CC peptide of the invention
 XX
 SQ Sequence 13 AA;
 Query Match 78.6%; Score 44; DB 6; Length 13;
 Best Local Similarity 55.6%; Pred. No. 5.7;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 GYWLTIWGX 10
 |||||

Db 1 GYWCNVWGL 9

RESULT 4
ADJ50760
ID ADJ50760 standard; peptide; 13 AA.
XX AC ADJ50760;
XX DT 06-MAY-2004 (first entry)
XX DE Human serum albumin binding peptide, Seq ID No 297.
XX DE human serum albumin; HSA; serum; blood; tumour; human.
XX KW Homo sapiens.
XX OS
XX PN WO2003106493-A1.
XX PD 24-DEC-2003.
XX PF 16-JUN-2003; 2003WO-US018896.
XX PR 14-JUN-2002; 2002US-0388642P.
XX PA (DYAX-) DYAX CORP.
XX PI Sato AK, Dawson BM;
XX WPI; 2004-082161/08.
XX Evaluating sample comprising soluble serum protein by forming complex
PT comprising serum protein and physically associated compounds using
PT peptide ligand that specifically binds with proteins, which is separated
PT and evaluated.

XX PS Disclosure; SEQ ID NO 297; 191pp; English.
XX The invention relates to a method of evaluating sample by providing a
CC soluble serum protein (I), one or more compounds physically associated
CC with (I), and a (I)-binding agent that comprises a peptide that
CC specifically binds to (I), allowing the (I)-binding agent to bind to (I)
CC to form a complex including one or more compounds physically associated
CC with (I), separating the complex from one or more components of the
CC sample, and evaluating one or more of the physically associated
CC compounds. The sample comprises blood or serum, or is obtained from a
CC biopsy. The sample may also be obtained from a tumour or a region within
CC 5 mm of a tumour. The method is useful for detecting modulators that
CC modulate interaction of serum protein-binding compound and serum protein
CC and for identifying binding ligands for serum protein. The present
CC sequence represents a serum albumin-binding peptide identified using the
CC method of the invention.

XX SQ Sequence 13 AA;
Query Match 78.6%; Score 44; DB 8; Length 13;
Best Local Similarity 55.6%; Pred. No. 5.7;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYWLTIWGX 10
||| :||:
Db 1 GYWCNVWGL 9

RESULT 5
AAB62747
ID AAB62747 standard; protein; 120 AA.
XX AC AAB62747;
XX DT 03-APR-2001 (first entry)
XX DE Human HIV-1 monoclonal antibody SEQ ID NO: 46.

XX Human immunodeficiency virus-1; HIV-1; human monoclonal antibody;
KW envelope glycoprotein; gp120; diagnosis.
XX Homo sapiens.
XX PN WO200100678-A1.
XX PD 04-JAN-2001.
XX PF 23-JUN-2000; 2000WO-US017327.
XX PR 30-JUN-1999; 99US-0141701P.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Watkins BA, Reitz MS;
XX WPI; 2001-112438/12.
XX DR N-PSDB; AAF29048.
XX Novel human monoclonal antibody immunoreactive with human
PT immunodeficiency virus-1 glycoprotein gp120, useful for detecting HIV-1
PT in biological sample and providing passive immunotherapy to HIV-1
PT infected mammal.
XX Claim 1; Page 51-52; 81pp; English.
XX The present invention provides the protein and coding sequences for the
CC variable regions of human monoclonal antibodies which are immunoreactive
CC with human immunodeficiency virus-1 (HIV-1) envelope glycoprotein gp120.
CC These can be used in diagnosis and therapy of HIV-1 infection

XX SQ Sequence 120 AA;
Query Match 78.6%; Score 44; DB 4; Length 120;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:||||:|:
Db 103 RGYWVSWMGQ 112

RESULT 6
ABU30004
ID ABU30004 standard; protein; 474 AA.
XX AC ABU30004;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #15531.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Enterococcus faecium.
XX WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA33874.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 57928; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 474 AA;

Query Match 78.6%; Score 44; DB 6; Length 474;
 Best Local Similarity 75.0%; Pred. No. 2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGFWLTIW 8
 Db 93 WGYWLTWCW 100
 :|||||
 111 WGYWLTWCW 118

RESULT 7
 ADC97318
 ID ADC97318 standard; protein; 492 AA.
 XX
 AC ADC97318;
 DT 01-JAN-2004 (first entry)
 XX
 DE E. faecium protein sequence SEQ ID 6945.
 XX
 KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
 KW abdominal-pelvic infection.
 XX
 OS Enterococcus faecium.
 XX
 FN US6583275-B1.
 XX
 PD 24-JUN-2003.
 XX

PF 10-JUN-1998; 98US-00107532.
 XX
 PR 02-JUL-1997; 97US-0051571P.
 PR 14-MAY-1998; 98US-0085598P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Doucette-Stamm LA, Bush D;
 XX
 DR WPI; 2003-799836/75.
 DR N-PSDB; ADC93664.
 XX
 CC New isolated nucleic acid derived from *Enterococcus faecium* encoding an
 CC Enterococcus faecium polypeptide useful for detection, prevention and
 CC treatment of a pathological condition resulting from a bacterial
 CC infection.
 XX
 PS Example 1; SEQ ID NO 6945; 243pp; English.
 XX
 CC The invention relates to an isolated nucleic acid derived from
 CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
 CC one of 10 fully defined sequences given in the (for comprising 40
 CC sequential nucleotides chosen from any of the nucleic acids, its
 CC complement or sequences hybridising to it). Also included are a
 CC recombinant vector comprising the nucleic acid operably linked to
 CC transcription regulatory element, a cell comprising the vector and a
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
 CC The nucleic acids is useful for diagnosing pathological conditions
 CC resulting from *E. faecium* bacterial infection (e.g. urinary tract
 CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
 CC infection) and for screening drugs such as agonists and antagonists. The
 CC nucleic acid is useful for recombinant production of *Candida albicans* -
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions
 CC and vaccines containing the nucleic acid are useful for preventing or
 CC treating Enterococcus faecium infections. The present sequence represents
 CC one if the disclosed *E. faecium* proteins.
 XX
 SQ Sequence 492 AA;

Query Match 78.6%; Score 44; DB 7; Length 492;
 Best Local Similarity 75.0%; Pred. No. 2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGFWLTIW 8
 Db 111 WGYWLTWCW 118
 :|||||

RESULT 8
 AAE09365
 ID AAE09365 standard; protein; 1140 AA.
 XX
 AC AAE09365;
 DT 19-NOV-2001 (first entry)
 XX
 DE Human ATP-binding cassette transporter ABCC6, MRP6 truncated mutant.
 XX
 KW Human; prenatal diagnosis; dermal lesion; cardiovascular disease; MRP6;
 KW Multidrug resistance-associated protein 6; macular degeneration; ABCC6;
 KW ATP-binding cassette transporter; arterial insufficiency; chromosome 16;
 KW pseudoxanthoma elasticum; PXE; heritable disorder; retinal haemorrhage;
 KW mutant; mutein.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FN WO200162977-A2.
 XX
 PD 30-AUG-2001.
 XX
 PF 23-FEB-2001; 2001WO-US005741.


```

FT      /note= "encoded by ACT"
FT Misc-difference 1274
FT      /note= "encoded by TGC"
FT Misc-difference 1287
FT      /note= "encoded by TTC"
FT Misc-difference 1455
FT      /note= "encoded by CCC"
XX
XX WO9949735-A1.
XX
XX 07-OCT-1999.
XX
XX 26-MAR-1999; 99WO-US006644.
XX
XX 27-MAR-1998; 98US-0079759P.
XX 03-AUG-1998; 98US-0095153P.
XX
XX (FOXC-) FOX CHASE CANCER CENT.
XX
XX Kruh G, Lee K, Belinsky M, Bain L;
XX
XX WPI; 1999-610812/52.
XX N-PSDB; AAZ30081.
XX
XX New transporter gene useful for screening for anti-cancer drugs.
XX
XX Claim 39; Page 144-147; 153pp; English.
XX
XX The present sequence represents a human MPR-related ABC transporter
XX (MOAT) protein, designated MOAT-E. The protein comprises a multi-domain
XX structure including a tandem repeat of nucleotide binding folds appended
XX C-terminal to a hydrophobic domain, having Walker A and B ATP binding
XX sites and several potential membrane spanning domains. The MOAT nucleic
XX acids are useful for screening a test compound for inhibition of MOAT
XX mediated transport, indicated by restoration of anticancer drug
XX sensitivity, which in turn causes a reduction of transporter mediated
XX cellular efflux of anticancer agents. MOAT DNA or RNA may be used as
XX probes to detect the presence or expression of genes encoding MOAT
XX proteins. Anti-MOAT antibodies are useful for detecting and quantitating
XX MOAT proteins
XX
XX Sequence 1503 AA;
XX
XX      Query Match      78.6%; Score 44; DB 2; Length 1503;
XX      Best Local Similarity 50.0%; Pred. No. 6.1e+02;
XX      Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
XX
XX      QY      1 XGYWLTWGX 10
XX      Db      964 RGYWLSLWAD 973
XX
XX      RESULT 11
XX      AAE09361
XX      ID AAE09361 standard; protein; 1503 AA.
XX      AC AAE09361;
XX
XX      19-NOV-2001 (first entry)
XX
XX      Human ATP-binding cassette transporter ABCC6 (MRP6) protein.
XX
XX      Human; prenatal diagnosis; dermal lesion; cardiovascular disease; MRP6;
XX      Multidrug Resistance-associated protein 6; macular degeneration; ABCC6;
XX      ATP-binding cassette transporter; arterial insufficiency; chromosome 16;
XX      Pseudoxanthoma elasticum; PXE; heritable disorder; retinal haemorrhage.
XX
XX      Homo sapiens.
XX
XX      Key      Location/Qualifiers
XX      FH      656..679
XX      FT      /note= "Nucleotide binding fold-1 (NFB-1) region"
XX      FT      747..768

```

```

FT      /note= "Nucleotide binding fold-1 (NFB-1) region"
FT 775..784
FT      /note= "Nucleotide binding fold-1 (NFB-1) region"
FT 1292..1307
FT      /note= "Nucleotide binding fold-2 (NFB-2) region"
FT 1321..1327
FT      /note= "Nucleotide binding fold-2 (NFB-2) region"
FT 1403..1433
FT      /note= "Nucleotide binding fold-2 (NFB-2) region"
XX
XX WO200162977-A2.
XX
XX 30-AUG-2001.
XX
XX 23-FEB-2001; 2001WO-US005741.
XX
XX 23-FEB-2000; 2000US-0184269P.
XX
XX (PXEI-) PXE INT INC.
XX (UYHA-) UNIV HAWAII.
XX
XX Boyd CD, Csizsar K, Lesaux O, Urban Z, Terry S;
XX
XX WPI; 2001-536645/59.
XX N-PSDB; AAD16230, AAD16231.
XX
XX Screening presence of Pseudoxanthoma elasticum mutation useful for
XX identifying homozygotes, compound heterozygotes or carriers involves
XX determining presence of mutation in MRP6 (ABCC6) nucleic acid.
XX
XX Example 2; Fig 3; 163pp; English.
XX
XX The invention relates to methods and compositions for diagnosing and
XX treating Pseudoxanthoma elasticum (PXE) and PXE associated physiological
XX dysfunctions. The invention is useful for screening for the presence of a
XX PXE mutation. Mutations associated with PXE maps to the ATP-binding
XX cassette transporter ABCC6 (MRP6-Multidrug Resistance associated protein-
XX 6) gene locus on chromosome 16. ABCC6 (MRP6) gene encodes a 165 kDa
XX protein located in the plasma membrane containing 17 membrane-spanning
XX helices grouped into three transmembrane domains. PXE is inherited as an
XX autosomal recessive phenotype or appears as a sporadic phenotype. PXE is
XX a heritable disorder characterised by mineralisation of elastic fibers in
XX skin, arteries and the retina, that result in dermal lesions with
XX associated laxity and loss of elasticity, arterial insufficiency,
XX cardiovascular disease and retinal haemorrhages leading to macular
XX degeneration. The method is useful for screening a population of
XX individuals in order to identify individuals with one or more PXE
XX associated MRP6 alleles who are then provided with appropriate genetic
XX counselling in view of the PXE status. The methods are useful for
XX identifying homozygotes, compound heterozygotes or carriers and thus are
XX useful in the area of genetic testing, carrier detection and prenatal
XX diagnosis. The present sequence is human ATP-binding cassette (ABC)
XX transporter, ABCC6 (MRP6) protein belonging to sub-family "C". Since
XX ABCC6 protein is involved in drug-resistance it is also called Multidrug
XX Resistance associated protein 6 (MRP6)
XX
XX Sequence 1503 AA;
XX
XX      Query Match      78.6%; Score 44; DB 4; Length 1503;
XX      Best Local Similarity 50.0%; Pred. No. 6.1e+02;
XX      Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
XX
XX      QY      1 XGYWLTWGX 10
XX      Db      964 RGYWLSLWAD 973
XX
XX      RESULT 12
XX      AAE09370
XX      ID AAE09370 standard; protein; 1503 AA.
XX      AC AAE09370;
XX      XX

```


Query Match 78.6%; Score 44; DB 4; Length 1503;
 Best Local Similarity 50.0%; Pred. NO. 6.1e+02;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
 :||||: :
 Db 964 RGYWLSLWAD 973

RESULT 14
 AAE09368
 ID AAE09368 standard; protein; 1503 AA.
 XX
 AC AAE09368;
 XX
 DT 19-NOV-2001 (first entry)
 XX
 DE Human ATP-binding cassette transporter ABCC6 (MRP6) G1302R mutant.
 XX
 KW Human; prenatal diagnosis; dermal lesion; cardiovascular disease; MRP6;
 KW Multidrug Resistance-associated protein 6; macular degeneration; ABCC6;
 KW ATP-binding cassette transporter; arterial insufficiency; Chromosome 16;
 KW Pseudoxanthoma elasticum; PXE; heritable disorder; retinal haemorrhage;
 KW mutant; mutain.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1302 /note= "Wild type Gly substituted with Arg"
 FT
 XX WO200162977-A2.
 XX
 XX 30-AUG-2001.
 XX
 XX 23-FEB-2001; 2001WO-US005741.
 XX
 XX 23-FEB-2000; 2000US-0184269P.
 XX
 XX (PXEI-) PXE INT INC.
 XX (UYHA-) UNIV HAWAII.
 XX
 XX Boyd CD, Csizsar K, Lesaux O, Urban Z, Terry S;
 XX WPI; 2001-536645/59.
 XX N-PSDB; AAD16261.
 XX
 XX Screening presence of Pseudoxanthoma elasticum mutation useful for
 XX identifying homozygotes, compound heterozygotes or carriers involves
 XX determining presence of mutation in MRP6 (ABCC6) nucleic acid.
 XX
 XX Claim 13; Page; 163pp; English.
 XX
 XX The invention relates to methods and compositions for diagnosing and
 XX treating Pseudoxanthoma elasticum (PXE) and PXE associated physiological
 XX dysfunctions. The invention is useful for screening for the presence of a
 XX PXE mutation. Mutations associated with PXE maps to the ATP-binding
 XX cassette transporter ABCC6 (MRP6-Multidrug Resistance associated protein-
 XX 6) gene locus on chromosome 16. ABCC6 (MRP6) gene encodes a 165 kDa
 XX protein located in the plasma membrane containing 17 membrane- spanning
 XX helices grouped into three transmembrane domains. PXE is inherited as an
 XX autosomal recessive phenotype or appears as a sporadic phenotype. PXE is
 XX a heritable disorder characterised by mineralisation of elastic fibers in
 XX skin, arteries and the retina, that result in dermal lesions with
 XX associated laxity and loss of elasticity, arterial insufficiency,
 XX cardiovascular disease and retinal haemorrhages leading to macular
 XX degeneration. The method is useful for screening a population of
 XX individuals in order to identify individuals with one or more PXE
 XX associated MRP6 alleles who are then provided with appropriate genetic
 XX counselling in view of the PXE status. The methods are useful for
 XX identifying homozygotes, compound heterozygotes or carriers and thus are

CC useful in the area of genetic testing, carrier detection and prenatal
 CC diagnosis. The present sequence is human ATP-binding cassette (ABC)
 CC transporter, ABCC6 (MRP6) G1302R mutant protein. Note: The present
 CC sequence is not shown in the specification but is derived from human ATP-
 CC binding cassette transporter ABCC6 (MRP6) protein [SED ID NO: 3] shown in
 CC figure 3 of the specification (AAE09361)
 XX
 SQ Sequence 1503 AA;
 Query Match 78.6%; Score 44; DB 4; Length 1503;
 Best Local Similarity 50.0%; Pred. NO. 6.1e+02;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
 :||||: :
 Db 964 RGYWLSLWAD 973

RESULT 15
 AAE09369
 ID AAE09369 standard; protein; 1503 AA.
 XX
 AC AAE09369;
 XX
 DT 19-NOV-2001 (first entry)
 XX
 DE Human ATP-binding cassette transporter ABCC6 (MRP6) A1303P mutant.
 XX
 KW Human; prenatal diagnosis; dermal lesion; cardiovascular disease; MRP6;
 KW Multidrug Resistance-associated protein 6; macular degeneration; ABCC6;
 KW ATP-binding cassette transporter; arterial insufficiency; chromosome 16;
 KW Pseudoxanthoma elasticum; PXE; heritable disorder; retinal haemorrhage;
 KW mutant; mutain.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1303 /note= "Wild type Ala substituted with Pro"
 FT
 XX WO200162977-A2.
 XX
 XX 30-AUG-2001.
 XX
 XX 23-FEB-2001; 2001WO-US005741.
 XX
 XX 23-FEB-2000; 2000US-0184269P.
 XX
 XX (PXEI-) PXE INT INC.
 XX (UYHA-) UNIV HAWAII.
 XX
 XX Boyd CD, Csizsar K, Lesaux O, Urban Z, Terry S;
 XX WPI; 2001-536645/59.
 XX N-PSDB; AAD16262.
 XX
 XX Screening presence of Pseudoxanthoma elasticum mutation useful for
 XX identifying homozygotes, compound heterozygotes or carriers involves
 XX determining presence of mutation in MRP6 (ABCC6) nucleic acid.
 XX
 XX Claim 13; Page; 163pp; English.
 XX
 XX The invention relates to methods and compositions for diagnosing and
 XX treating Pseudoxanthoma elasticum (PXE) and PXE associated physiological
 XX dysfunctions. The invention is useful for screening for the presence of a
 XX PXE mutation. Mutations associated with PXE maps to the ATP-binding
 XX cassette transporter ABCC6 (MRP6-Multidrug Resistance associated protein-
 XX 6) gene locus on chromosome 16. ABCC6 (MRP6) gene encodes a 165 kDa
 XX protein located in the plasma membrane containing 17 membrane- spanning
 XX helices grouped into three transmembrane domains. PXE is inherited as an
 XX autosomal recessive phenotype or appears as a sporadic phenotype. PXE is
 XX a heritable disorder characterised by mineralisation of elastic fibers in
 XX skin, arteries and the retina, that result in dermal lesions with
 XX associated laxity and loss of elasticity, arterial insufficiency,
 XX cardiovascular disease and retinal haemorrhages leading to macular
 XX degeneration. The method is useful for screening a population of
 XX individuals in order to identify individuals with one or more PXE
 XX associated MRP6 alleles who are then provided with appropriate genetic
 XX counselling in view of the PXE status. The methods are useful for
 XX identifying homozygotes, compound heterozygotes or carriers and thus are

CC skin, arteries and the retina, that result in dermal lesions with
 CC associated laxity and loss of elasticity, arterial insufficiency,
 CC cardiovascular disease and retinal haemorrhages leading to macular
 CC degeneration. The method is useful for screening a population of
 CC individuals in order to identify individuals with one or more PXE
 CC associated MRP6 alleles who are then provided with appropriate genetic
 CC counselling in view of the PXE status. The methods are useful for
 CC identifying homozygotes, compound heterozygotes or carriers and thus are
 CC useful in the area of genetic testing, carrier detection and prenatal
 CC diagnosis. The present sequence is human ATP-binding cassette (ABC)
 CC transporter, ABC6 (MRP6) A1303P mutant protein. Note: The present
 CC sequence is not shown in the specification but is derived from human ATP-
 CC binding cassette transporter ABC6 (MRP6) protein [SED ID NO: 3] shown in
 CC figure 3 of the specification (AAE09361)
 XX
 SQ Sequence 1503 AA;

Query Match 78.6%; Score 44; DB 4; Length 1503;
 Best Local Similarity 50.0%; Pred. No. 6.1e+02;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
 Db :|||||: :
 964 RGIWLSLWAD 973

Search completed: January 3, 2005, 16:27:20
 Job time : 114.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:19:47 ; Search time 28.8 Seconds
(without alignments)
23.027 Million cell updates/sec

Title: US-10-046-922-34
Perfect score: 56
Sequence: 1 XGYWLTWGX 10

Scoring table: BLOSUM62DX*

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata1/iaa/5A COMB.pdp.*
2: /cgn2_6/ptodata1/iaa/5B COMB.pdp.*
3: /cgn2_6/ptodata1/iaa/6A COMB.pdp.*
4: /cgn2_6/ptodata1/iaa/6B COMB.pdp.*
5: /cgn2_6/ptodata1/iaa/PTUS COMB.pdp.*
6: /cgn2_6/ptodata1/iaa/backfiles1.pdp.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	78.6	492	4 US-09-107-532A-6945	Sequence 6945, Ap
2	44	78.6	1498	4 US-09-792-616-9	Sequence 9, Appli
3	44	78.6	1503	4 US-09-792-616-3	Sequence 3, Appli
4	43	76.8	222	4 US-09-071-035-408	Sequence 408, App
5	43	76.8	229	4 US-09-134-000C-3630	Sequence 3630, Ap
6	43	76.8	266	4 US-09-071-035-406	Sequence 406, App
7	41	73.2	668	4 US-09-248-796A-19350	Sequence 19350, A
8	40	71.4	328	4 US-09-540-236-2629	Sequence 2629, Ap
9	40	71.4	478	4 US-09-107-532A-6868	Sequence 6868, Ap
10	38.5	68.8	612	4 US-09-252-991A-19134	Sequence 19134, A
11	38	67.9	24	4 US-09-270-767-41173	Sequence 41173, A
12	38	67.9	24	4 US-09-270-767-56389	Sequence 56389, A
13	38	67.9	119	2 US-08-318-157B-5	Sequence 5, Appli
14	38	67.9	119	4 US-09-253-794-5	Sequence 5, Appli
15	38	67.9	132	4 US-09-248-796A-22785	Sequence 22785, A
16	38	67.9	278	4 US-09-145-828A-11	Sequence 11, Appl
17	38	67.9	278	4 US-09-903-456-18	Sequence 18, Appl
18	38	67.9	499	4 US-09-252-991A-23328	Sequence 23328, A
19	38	67.9	591	4 US-09-561-077C-18	Sequence 18, Appl
20	38	67.9	591	4 US-09-221-014-18	Sequence 18, Appl
21	38	67.9	865	4 US-09-252-991A-19339	Sequence 19339, A
22	37	66.1	172	4 US-09-107-532A-6560	Sequence 6560, Ap
23	37	66.1	219	3 US-09-247-373B-52	Sequence 52, Appl
24	37	66.1	227	4 US-09-270-767-32789	Sequence 32789, A
25	37	66.1	227	4 US-09-270-767-48006	Sequence 48006, A
26	37	66.1	242	2 US-08-512-955-4	Sequence 4, Appli
27	37	66.1	247	4 US-09-252-991A-26736	Sequence 26736, A

28	37	66.1	263	4 US-09-134-000C-4830	Sequence 4830, Ap
29	37	66.1	265	4 US-09-489-039A-8793	Sequence 8793, Ap
30	37	66.1	268	2 US-08-861-512-3	Sequence 3, Appli
31	37	66.1	280	1 US-08-414-685-2	Sequence 2, Appli
32	37	66.1	305	4 US-09-252-991A-27334	Sequence 27334, A
33	37	66.1	367	4 US-09-248-796A-15188	Sequence 15188, A
34	37	66.1	1528	1 US-08-463-092B-6	Sequence 6, Appli
35	37	66.1	1528	2 US-08-462-109A-6	Sequence 6, Appli
36	37	66.1	1528	3 US-08-460-907B-6	Sequence 6, Appli
37	37	66.1	1528	3 US-08-463-179A-6	Sequence 6, Appli
38	37	66.1	1528	3 US-08-461-384B-6	Sequence 6, Appli
39	37	66.1	1531	1 US-08-141-893-2	Sequence 2, Appli
40	37	66.1	1531	1 US-08-463-092B-2	Sequence 2, Appli
41	37	66.1	1531	1 US-08-463-092B-4	Sequence 4, Appli
42	37	66.1	1531	2 US-08-462-109A-2	Sequence 2, Appli
43	37	66.1	1531	2 US-08-462-109A-4	Sequence 4, Appli
44	37	66.1	1531	2 US-08-460-907B-2	Sequence 2, Appli
45	37	66.1	1531	2 US-08-460-907B-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-107-532A-6945
; Sequence 6945, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinieglo, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6945:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 492 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...492
; SEQUENCE DESCRIPTION: SEQ ID NO: 6945:
US-09-107-532A-6945

Query Match 78.6%; Score 44; DB 4; Length 492;
Best Local Similarity 75.0%; Pred. No. 75;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTW 8
:|||||
Db 111 WGYWLTW 118

RESULT 2
US-09-792-616-9
; Sequence 9, Application US/09792616
; Patent No. 6780587
; GENERAL INFORMATION:
; APPLICANT: PXE International, Inc.
; APPLICANT: University of Hawaii
; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (MRP6) causing
; TITLE OF INVENTION: Pseudoxanthoma Elasticum
; FILE REFERENCE: PXE-001
; CURRENT APPLICATION NUMBER: US/09/792,616
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1498
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-792-616-9

Query Match 78.6%; Score 44; DB 4; Length 1498;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:|||||:
Db 959 QGYWLSLWAD 968

RESULT 3
US-09-792-616-3
; Sequence 3, Application US/09792616
; Patent No. 6780587
; GENERAL INFORMATION:
; APPLICANT: PXE International, Inc.
; APPLICANT: University of Hawaii
; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (MRP6) causing
; TITLE OF INVENTION: Pseudoxanthoma Elasticum
; FILE REFERENCE: PXE-001
; CURRENT APPLICATION NUMBER: US/09/792,616
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1503
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-616-3

Query Match 78.6%; Score 44; DB 4; Length 1503;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:|||||:
Db 964 RGYWLSLWAD 973

RESULT 4
US-09-071-035-408
; Sequence 408, Application US/090710135
; Patent No. 6448043
; GENERAL INFORMATION:

APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB169P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 408:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-408

Query Match 76.8%; Score 43; DB 4; Length 222;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:|||||:
Db 178 AGTWILWQ 187

RESULT 5
US-09-134-000C-3630
; Sequence 3630, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3630
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3630

Query Match 76.8%; Score 43; DB 4; Length 229;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10

db 202 AGTWITLWQ 211

RESULT 6

US-09-071-035-406
Sequence 406, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

REF: 20830
COMPUTER READABLE FORM:

```

1 MEDIUM TYPE: Diskette, 3.50 inch, 1
2
3 COMPUTER: HP Vectra 486/33
4
5 OPERATING SYSTEM: MSDOS version 6.2
6
7 SOFTWARE: ASCII Text
8
9 CURRENT APPLICATION DATA:
10 APPLICATION NUMBER: US/09/071,035
11 FILING DATE:
12
13 CLASSIFICATION:
14
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER:
17 FILING DATE:
18
19 ATTORNEY/AGENT INFORMATION:
20 NAME: A. Anders Brookes
21
22 REGISTRATION NUMBER: 36,373
23
24 REFERENCE/DOCKET NUMBER: PB369P2
25
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: (301) 309-8504
28
29 TELEFAX: (301) 309-8512
30
31 INFORMATION FOR SEQ ID NO: 406:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 266 amino acids
34 TYPE: amino acid
35 STRANDEDNESS: single
36 TOPOLOGY: linear
37
38 MOLECULE TYPE: protein
39
40 US-09-071-035-406

```

Query Match 76.8%; Score 43; DB 4; Length 266;
Best Local Similarity 50.0%; Pred. No. 56;
Matches 5: Conservative 4; Mismatches 1: Indels

```
Qy      1 XGYWLTWGX 10
        :| |:|:|:|:
Db     202 AGTWITLWGO 211
```

RESULT, T 7

RESOLUTION 7
US-09-248-796A-19350
; Sequence 19350, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19350

; LENGTH: 668
 ; TYPE: PRT
 ; ORGANISM: C
 US-09-248-796A-

Query Match 73.2%; Score 41; DB 4; Length 668;
Best Local Similarity 62.5%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels

Qy 3 YWLTWGX 10
|||:||||
Db 372 YWMTIWNA 379

RESULTS

```

US-09-540-236-2629
; Sequence 2629, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND
; TITLE OF INVENTION: FOR DIAGNOSTICS A
; FILE REFERENCE: 2709,2005-001
; CURRENT APPLICATION NUMBER: US/09/540
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2629
; LENGTH: 328
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2629

```

Query Match 71.4%; Score 40; DB 4; Length 328;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 5: Conservative 3; Mismatches 2; Indels

Qy 1 XGYWLTWGX 10
:||||:
db 115 TGOWLTWSG 124

RESIT.T 9

US-09-107-532A-6868
Sequence 6868, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: LYNN A Douchette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489

```
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6868:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...478
; SEQUENCE DESCRIPTION: SEQ ID NO: 6868:
US-09-107-532A-6868

Query Match 71.4%; Score 40; DB 4; Length 478;
Best Local Similarity 62.5%; Pred. No. 2.7e+02; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWIM 8
   :||||:|
Db 97 WGYWLSAW 104

RESULT 10
US-09-252-991A-19134
; Sequence 19134, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19134
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19134

Query Match 68.8%; Score 38.5; DB 4; Length 612;
Best Local Similarity 54.5%; Pred. No. 5.6e+02; Indels 1; Gaps 1;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 XGYW-LTIWGX 10
   :|||||:|
Db 49 VGYWLSLEWVG 59

RESULT 11
US-09-270-767-41173
; Sequence 41173, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41173
; LENGTH: 24

; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6868:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...478
; SEQUENCE DESCRIPTION: SEQ ID NO: 6868:
US-09-107-532A-6868

Query Match 71.4%; Score 40; DB 4; Length 478;
Best Local Similarity 62.5%; Pred. No. 2.7e+02; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWIM 8
   :||||:|
Db 97 WGYWLSAW 104

RESULT 10
US-09-252-991A-19134
; Sequence 19134, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19134
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19134

Query Match 68.8%; Score 38.5; DB 4; Length 612;
Best Local Similarity 54.5%; Pred. No. 5.6e+02; Indels 1; Gaps 1;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 XGYW-LTIWGX 10
   :|||||:|
Db 49 VGYWLSLEWVG 59

RESULT 11
US-09-270-767-41173
; Sequence 41173, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41173
; LENGTH: 24
```

```
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-41173

Query Match 67.9%; Score 38; DB 4; Length 24;
Best Local Similarity 50.0%; Pred. No. 26; Indels 0; Gaps 0;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 YWLTWGX 10
   :|||:|
Db 16 HWLSLWGL 23

RESULT 12
US-09-270-767-56389
; Sequence 56389, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56389
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-56389

Query Match 67.9%; Score 38; DB 4; Length 24;
Best Local Similarity 50.0%; Pred. No. 26; Indels 0; Gaps 0;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 YWLTWGX 10
   :|||:|
Db 16 HWLSLWGL 23

RESULT 13
US-08-318-157B-5
; Sequence 5, Application US/08318157B
; Patent No. 5874540
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; APPLICANT: ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
```

TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-318-157B-5

Query Match 67.9%; Score 38; DB 2; Length 119;
Best Local Similarity 40.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
DB 102 AGCWIDVWGQ 111

RESULT 14

US-09-253-794-5
Sequence 5, Application US/09253794
Patent No. 6676924
GENERAL INFORMATION:
APPLICANT: HANSEN, Hans J.
ARMOUR, Kathryn L.
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
MOUSE MONOCLONAL ANTIBODIES

NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/253,794
FILING DATE: 22-Feb-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,157
FILING DATE: 05-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: SAXE, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 18733/464
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-253-794-5

Query Match 67.9%; Score 38; DB 4; Length 119;
Best Local Similarity 40.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
DB 102 AGCWIDVWGQ 111

RESULT 15

US-09-248-796A-22785
Sequence 22785, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 22785
LENGTH: 132
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: UNSURE
LOCATION: (6),(77),(85)
OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown

US-09-248-796A-22785

Query Match 67.9%; Score 38; DB 4; Length 132;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYWLT 7
DB 122 NGYWLTI 128

Search completed: January 3, 2005, 16:33:48
Job time : 29.8 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:32:45 ; Search time 100.8 Seconds
(without alignments)
35.687 Million cell updates/sec

Title: US-10-046-922-34

Perfect score: 56

Sequence: 1 XGYWLTIGX 10

Scoring table: BLOSUM62DX
Gapop 10.0, Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	10	13	US-10-046-922-34
2	56	100.0	10	13	US-10-046-922-35
3	44	78.6	13	14	US-10-125-869A-73
4	44	78.6	13	15	US-10-462-282-297
5	44	78.6	474	15	US-10-282-122A-57928
6	44	78.6	671	16	US-10-437-963-190740
7	44	78.6	1498	16	US-09-792-616-9
8	44	78.6	1498	16	US-10-764-328-9
9	44	78.6	1503	10	US-09-792-616-3
10	44	78.6	1503	16	US-10-764-328-3
11	43	76.8	34	9	US-09-864-761-43458
12	43	76.8	222	9	US-09-071-035-408
13	43	76.8	222	14	US-10-206-576-408

14	43	76.8	266	9	US-09-071-035-406	Sequence 406, App
15	43	76.8	266	14	US-10-206-576-406	Sequence 406, App
16	42	75.0	14	14	US-10-125-869A-114	Sequence 114, App
17	42	75.0	14	15	US-10-462-282-338	Sequence 338, App
18	42	75.0	82	15	US-10-424-599-240891	Sequence 240891, App
19	42	75.0	116	10	US-09-910-483-1	Sequence 1, Appli
20	42	75.0	116	10	US-09-910-483-5	Sequence 5, Appli
21	42	75.0	116	10	US-09-910-483-9	Sequence 9, Appli
22	42	75.0	116	10	US-09-910-483-13	Sequence 13, Appli
23	42	75.0	116	10	US-09-910-483-17	Sequence 17, Appli
24	42	75.0	116	10	US-09-910-483-21	Sequence 21, Appli
25	42	75.0	116	10	US-09-910-483-25	Sequence 25, Appli
26	42	75.0	116	10	US-09-910-483-29	Sequence 29, Appli
27	42	75.0	116	10	US-09-910-483-33	Sequence 33, Appli
28	42	75.0	116	10	US-09-910-483-37	Sequence 37, Appli
29	42	75.0	116	10	US-09-910-483-41	Sequence 41, Appli
30	42	75.0	116	10	US-09-910-483-43	Sequence 43, Appli
31	42	75.0	138	14	US-10-160-232-86	Sequence 86, Appli
32	42	75.0	138	14	US-10-160-232-90	Sequence 90, Appli
33	41	73.2	373	14	US-10-369-493-8834	Sequence 8834, Ap
34	40.5	72.3	101	17	US-10-425-115-260407	Sequence 260407,
35	40	71.4	13	14	US-10-125-869A-75	Sequence 75, Appli
36	40	71.4	13	15	US-10-462-282-299	Sequence 299, App
37	40	71.4	25	14	US-10-280-066-476	Sequence 476, App
38	40	71.4	43	9	US-09-864-761-47847	Sequence 47847, A
39	40	71.4	57	15	US-10-424-599-226815	Sequence 226815,
40	40	71.4	86	17	US-10-425-115-188067	Sequence 188067,
41	40	71.4	91	14	US-10-156-761-13764	Sequence 13764, A
42	40	71.4	101	15	US-10-424-599-194869	Sequence 194869,
43	40	71.4	162	9	US-09-738-626-5178	Sequence 5178, Ap
44	40	71.4	167	17	US-10-425-115-334071	Sequence 334071,
45	40	71.4	177	15	US-10-276-774-1797	Sequence 1797, Ap

ALIGNMENTS

RESULT 1
US-10-046-922-34
; Sequence 34, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Isolated peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(1)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (10)..(10)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-34

Query Match 100.0%; Score 56; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 XGYWLTIGX 10
DB 1 XGYWLTIGX 10

```
RESULT 2
US-10-046-922-35
; Sequence 35, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Ekki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35
; LENGTH: 10
; TYPE: PRT
; ORGANISM: isolated peptide
US-10-046-922-35

Query Match      100.0%; Score 56; DB 13; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.18;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 XGYWLTWGX 10
DB      1 CGYWLTWGC 10

RESULT 3
US-10-125-869A-73
; Sequence 73, Application US/10125869A
; Publication No. US20030199671A1
; GENERAL INFORMATION:
; APPLICANT: Rondon, Isaac Jesus
; APPLICANT: Wu, Qi-Long
; APPLICANT: Ley, Arthur C.
; APPLICANT: Stochl, Mark
; APPLICANT: Ransohoff, Thomas C.
; APPLICANT: Potter, M. Daniel (deceased)
; TITLE OF INVENTION: BINDING MOLECULES FOR FC-REGION
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 3421.1006-001
; CURRENT APPLICATION NUMBER: US/10/125,869A
; CURRENT FILING DATE: 2002-11-19
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Fc region binding polypeptide
US-10-125-869A-73

Query Match      78.6%; Score 44; DB 14; Length 13;
Best Local Similarity 55.6%; Pred. No. 11;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GYWLTIWGX 10
DB      1 GYWCNVWGL 9

RESULT 4
US-10-462-262-297
; Sequence 297, Application US/10462262
; Publication No. US20040009534A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Aaron K.
; APPLICANT: Dawson, Bruce M.
; APPLICANT:

; TITLE OF INVENTION: PROTEIN ANALYSIS
; FILE REFERENCE: 10280-052001
; CURRENT APPLICATION NUMBER: US/10/462,262
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/388,642
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 297
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: immunoglobulin binding polypeptide
US-10-462-262-297

Query Match      78.6%; Score 44; DB 15; Length 13;
Best Local Similarity 55.6%; Pred. No. 11;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GYWLTIWGX 10
DB      1 GYWCNVWGL 9

RESULT 5
US-10-282-122A-57928
; Sequence 57928, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57928
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Enterococcus faecium
```

US-10-282-122A-57928

Query Match 78.6%; Score 44; DB 15; Length 474;
Best Local Similarity 75.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTW 8

DB 93 WGYWLTW 100

RESULT 6

US-10-437-963-190740

; Sequence 190740, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 190740

; LENGTH: 671

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT MRT4530_87125C.1.pep

US-10-437-963-190740

Query Match

Best Local Similarity 78.6%; Score 44; DB 16; Length 671;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 YWLTWGX 10

DB 329 YWTTWGR 336

RESULT 7

US-09-792-616-9

; Sequence 9, Application US/09792616

; Publication No. US20030165828A1

; GENERAL INFORMATION:

; APPLICANT: PXE International, Inc.

; APPLICANT: University of Hawaii

; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (MRP6) causing

; TITLE OF INVENTION: Pseudoxanthoma Elasticum

; FILE REFERENCE: PXE-001

; CURRENT APPLICATION NUMBER: US/09/792,616

; CURRENT FILING DATE: 2001-02-23

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 9

; LENGTH: 1498

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-792-616-9

Query Match

Best Local Similarity 78.6%; Score 44; DB 10; Length 1498;

Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10

:|||||:

DB 959 QGYWLSLWAD 968

RESULT 8

US-10-764-328-9

; Sequence 9, Application US/10764328

; Publication No. US20040166521A1

; GENERAL INFORMATION:

; APPLICANT: PXE International, Inc.

; APPLICANT: University of Hawaii

; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (MRP6) causing

; TITLE OF INVENTION: Pseudoxanthoma Elasticum

; FILE REFERENCE: PXE-001PC

; CURRENT APPLICATION NUMBER: US/10/764,328

; CURRENT FILING DATE: 2004-01-23

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 9

; LENGTH: 1498

; TYPE: PRT

; ORGANISM: Mus musculus

US-10-764-328-9

Query Match

Best Local Similarity 78.6%; Score 44; DB 16; Length 1498;

Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10

DB 959 QGYWLSLWAD 968

RESULT 9

US-09-792-616-3

; Sequence 3, Application US/09792616

; Publication No. US20030165828A1

; GENERAL INFORMATION:

; APPLICANT: PXE International, Inc.

; APPLICANT: University of Hawaii

; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (MRP6) causing

; TITLE OF INVENTION: Pseudoxanthoma Elasticum

; FILE REFERENCE: PXE-001

; CURRENT APPLICATION NUMBER: US/09/792,616

; CURRENT FILING DATE: 2001-02-23

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 3

; LENGTH: 1503

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-792-616-3

Query Match

Best Local Similarity 78.6%; Score 44; DB 10; Length 1503;

Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10

DB 964 RGYWLSLWAD 973

RESULT 10

US-10-764-328-3

; Sequence 3, Application US/10764328

; Publication No. US20040166521A1

; GENERAL INFORMATION:

; APPLICANT: PXE International, Inc.

; APPLICANT: University of Hawaii

; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (MRP6) causing

; TITLE OF INVENTION: Pseudoxanthoma Elasticum

; FILE REFERENCE: PXE-001PC

; CURRENT APPLICATION NUMBER: US/10/764,328

; CURRENT FILING DATE: 2004-01-23

; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1503
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-328-3

Query Match 78.6%; Score 44; DB 16; Length 1503;
Best Local Similarity 50.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:||||:|:
Db 964 RGYWLSLWAD 973

RESULT 11
US-09-864-761-43458
; Sequence 43458, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43458
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens

; FEATURE:
; OTHER INFORMATION: MAP TO AC000403.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
US-09-864-761-43458

Query Match 76.8%; Score 43; DB 9; Length 34;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:||||:|:
Db 24 HGYWITVFGD 33

RESULT 12
US-09-071-035-408
; Sequence 408, Application US/09071035
; Publication No. US20020045737A1
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; APPLICATION DATA:
; CURRENT APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 408:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-408

Query Match 76.8%; Score 43; DB 9; Length 222;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:||||:|:
Db 178 AGTWITLMGQ 187

RESULT 13
US-10-206-576-408


```
/ Sequence 408, Application US/10206576
/ Publication No. US20030017495A1
/ GENERAL INFORMATION:
/ APPLICANT: Choi et al.
/ TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
/ NUMBER OF SEQUENCES: 497
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD-R
/ COMPUTER: Dell Latitude
/ OPERATING SYSTEM: Windows 98
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/206,576
/ FILING DATE: 29-Jul-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 09/071,035
/ FILING DATE: 1998-05-04
/ APPLICATION NUMBER: US 60/046,655
/ FILING DATE: 1997-05-16
/ APPLICATION NUMBER: US 60/044,031
/ FILING DATE: 1997-05-06
/ APPLICATION NUMBER: US 60/066,009
/ FILING DATE: 1997-11-14
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hyman, Mark J.
/ REGISTRATION NUMBER: 46,789
/ REFERENCE/DOCKET NUMBER: PB369P1D1
/ INFORMATION FOR SEQ ID NO: 408:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 222 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 408:
US-10-206-576-408
Query Match 76.8%; Score 43; DB 14; Length 222;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
Db 178 AGTWITLWGX 187

RESULT 14
US-09-071-035-406
/ Sequence 406, Application US/09071035
/ Publication No. US20020045737A1
/ GENERAL INFORMATION:
/ APPLICANT: Gil H. Choi
/ TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
/ NUMBER OF SEQUENCES: 496
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
```

```
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/071,035
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: A. Anders Brookes
/ REGISTRATION NUMBER: 36,373
/ REFERENCE/DOCKET NUMBER: PB369P2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 406:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 266 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-09-071-035-406
Query Match 76.8%; Score 43; DB 9; Length 266;
Best Local Similarity 50.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
Db 202 AGTWITLWGX 211

RESULT 15
US-10-206-576-406
/ Sequence 406, Application US/10206576
/ Publication No. US20030017495A1
/ GENERAL INFORMATION:
/ APPLICANT: Choi et al.
/ TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
/ NUMBER OF SEQUENCES: 497
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD-R
/ COMPUTER: Dell Latitude
/ OPERATING SYSTEM: Windows 98
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/206,576
/ FILING DATE: 29-Jul-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 09/071,035
/ FILING DATE: 1998-05-04
/ APPLICATION NUMBER: US 60/046,655
/ FILING DATE: 1997-05-16
/ APPLICATION NUMBER: US 60/044,031
/ FILING DATE: 1997-05-06
/ APPLICATION NUMBER: US 60/066,009
/ FILING DATE: 1997-11-14
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hyman, Mark J.
/ REGISTRATION NUMBER: 46,789
/ REFERENCE/DOCKET NUMBER: PB369P1D1
/ INFORMATION FOR SEQ ID NO: 406:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 266 amino acids
```

```

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 406:
US-10-206-576-406

Query Match      76.8%; Score 43; DB 14; Length 266;
Best Local Similarity 50.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      1 XGYWLTWGX 10
Db      202 AGTWITLWQ 211
      :|:|:|:|:
      :|:|:|:|:

Search completed: January 3, 2005, 16:54:14
Job time : 101.8 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:20:13 ; Search time 22.4 Seconds
(without alignments)
42.954 Million cell updates/sec

Title: US-10-046-922-34

Perfect score: 56

Sequence: 1 XGYWLTWGX 10

Scoring table: BLOSUM62DX, Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	78.6	441	2 C95307	probable transport
2	44	78.6	1502	2 T42216	multidrug resistant
3	43	76.8	419	2 E90446	permease [imported
4	42	75.0	142	2 C34903	replication chain pre
5	42	75.0	391	2 C74117	probable Na ⁺ /H ⁺ -ex
6	42	75.0	425	2 B71038	hypothetical prote
7	41	73.2	376	2 AF1978	arginine/ornithine
8	41	73.2	490	2 C86879	probable ABC trans
9	41	73.2	508	2 C95282	na ⁺ /h ⁺ antiporter
10	40	71.4	426	2 C75176	arginine/ornithine
11	40	71.4	475	2 T46745	probable oxidoredu
12	40	71.4	583	2 T50103	ins P4-binding pro
13	40	71.4	829	2 S58888	ins P4-binding pro
14	40	71.4	829	2 S71847	cation efflux syst
15	39.5	70.5	1063	2 A33830	cadmium, zinc, cob
16	39.5	70.5	1063	2 JCA700	probable integral
17	39	69.6	118	2 T35739	probable exported
18	39	69.6	167	2 AF0881	ubiquinol-cytochro
19	39	69.6	187	1 A13868	hypothetical prote
20	39	69.6	218	2 S76385	rod shape-determ
21	39	69.6	416	2 AF1127	ABC transporter [im
22	39	69.6	477	2 G90251	probable transcarb
23	38.5	68.8	607	2 F82966	ig heavy chain v r
24	38	67.9	123	1 AVMS14	hypothetical prote
25	38	67.9	172	2 A12332	hypothetical prote
26	38	67.9	250	2 A65843	protein F41H10.7 [
27	38	67.9	286	2 E88690	LysR-type protein
28	38	67.9	292	2 S49164	CDP-glucose 4,6-de
29	38	67.9	360	2 S74751	

RESULT 1

C95307

probable transport protein Sma0684 [imported] - Sinorhizobium meliloti (strain 1021) meg
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: C95307
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowse
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.,
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9886, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: C95307
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-441 <XUR>

A:Cross-references: UNIPROT:Q92ZT6; GB:AE006469; PIDN:AAK65021.1; PID:gl4523451; GSPDB:G

A:Experimental source: strain 1021, megaplasmid pSyma

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, P.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.,

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: Sma0684

A:Genome: plasmid

C:Superfamily: L-lysine transport protein

Query Match 78.6%; Score 44; DB 2; Length 441;

Best Local Similarity 50.0%; Pred. No. 25;

Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10

Db 91 WGYWISWAS 100

RESULT 2

T42216

multidrug resistance-associated protein homolog MLP-1 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T42216

R:Hirohashi, T.; Suzuki, H.; Ito, K.; Ogawa, K.; Kume, K.; Shimizu, T.; Sugiyama, Y.

Mol. Pharmacol. 53, 1068-1075, 1998

A:Title: Hepatic expression of multidrug resistance-associated protein-like proteins main

A:Reference number: Z22081; MUID:98279126; PMID:9614210

A:Accession: T42216

A:Status: preliminary; translated from GB/EMBL/DBJ

ig heavy chain pre
alpha-ketoglutarat
probable amino aci
arginine/ornithine
arginine/ornithine
YSD83 protein - ye
hypothetical prote
phosphatidylglycer
hypothetical prote
hypothetical prote
1-acyl-sn-glycerol
1-acylglycerol-3-p
probable mcfp prot
GP12 protein - ye
hypothetical prote
glycosyl transfera

ALIGNMENTS

380 2 S12839

435 2 H63607

472 2 E83497

482 2 JH0110

497 2 G86878

385 2 S43540

164 2 F69990

171 2 E83140

197 2 AF2356

241 2 G82391

247 2 G81013

247 2 A81957

257 2 E75325

280 2 S61111

292 2 A83779

294 2 D82657

A:Molecule type: mRNA
A:Residues: 1-1502 <HIR>
A:Cross-references: UNIPROT:O88269; EMBL:AB010466; NID:g3242457; PIDN:BA28954.1.; PID:g3242457
A:Experimental source: Strain Sprague-Dawley; liver
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 78.6%; Score 44; DB 2; Length 1502;
Best Local Similarity 50.0%; Pred. No. 80;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:|||||:
Db 963 QGYWLSLWAD 972

RESULT 3
E90446
permease [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: E90446
R;She, Q.; Singh, R.K.; Confalonieri, P.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: E90446
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-419 <KUR>
A:Cross-references: UNIPROT:Q97VB7; GB:AE006641; NID:g13916037; PIDN:AAK42828.1.; GSPDB:G
C:Genetics:
A:Gene: SS02718

Query Match 76.8%; Score 43; DB 2; Length 419;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:|||||:
Db 364 IGFWETLWGI 373

RESULT 4
C34903
Ig heavy chain precursor V region (5-27) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 16-Aug-1996
C:Accession: C34903
R;Bedzyk, W.D.; Herron, J.N.; Edmundson, A.B.; Voss Jr., E.W.
J. Biol. Chem. 265, 133-138, 1990
A:Title: Active site structure and antigen binding properties of idiotypically cross-rea
A:Reference number: A34903; MUID:90094387; PMID:2104617
A:Accession: C34903
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-142 <BED>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F;34-119/Domain: immunoglobulin homology <IMM>

Query Match 75.0%; Score 42; DB 2; Length 142;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:|||||:
Db 125 GGIWFAYWGO 134

RESULT 5
PC4117

replication protein homolog - Pyrococcus sp. (fragment)
N:Alternate names: hypothetical 391 protein
C:Species: Pyrococcus sp.
C:Date: 07-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 17-May-1996
C:Accession: PC4117
R;Rashid, N.; Morikawa, M.; Imanaka, T.
Gene 166, 139-143, 1995
A:Title: An abnormally acidic TATA-binding protein from a hyperthermophilic archaeon.
A:Reference number: JC4514; MUID:196105215; PMID:8529878
A:Accession: PC4117
A:Molecule type: DNA
A:Residues: 1-391 <RAS>
A:Cross-references: DDBJ:D50018

Query Match 75.0%; Score 42; DB 2; Length 391;
Best Local Similarity 75.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 YWLTWGX 10
:|||||:
Db 155 YWLTWGX 162

RESULT 6
B71038
probable Na+/H+-exchanging protein - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C:Accession: B71038
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi,
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic ar
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: B71038
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-425 <KAW>
A:Cross-references: UNIPROT:O59255; GB:AP000006; NID:g3236133; PIDN:BAA30706.1.; PID:g325
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1594

Query Match 75.0%; Score 42; DB 2; Length 425;
Best Local Similarity 60.0%; Pred. No. 48;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:|||||:
Db 318 FGAWLTWAGR 327

RESULT 7
AF1978
hypothetical protein alr1377 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AF1978
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimura, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF1978
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <KUR>
A:Cross-references: UNIPROT:Q8YX40; GB:BA000019; PIDN:BA073334.1.; PID:g17130724; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr1377

Query Match 73.2%; Score 41; DB 2; Length 376;
 Best Local Similarity 60.0%; Pred. No. 60;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
 : ||| |||
 Db 123 DGEWLRINGY 132

RESULT 8
 C86879
 A:Description: arginine/ornithine antiporter [imported] - Lactococcus lactis subsp. lactis (strain IL1404)
 C:Species: Lactococcus lactis subsp. lactis
 C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
 A:Accession: C86879
 R:Bolotin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
 Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp.
 A:Reference number: A86625; MUID:21235186; PMID:11337471
 A:Accession: C86879
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-490 <STO>
 A:Cross-references: UNIPROT:Q9CEI5; GB:AE005176; PID:g12725084; PIDN:AAK06133.1; GSPDB:G
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: arcD1
 C:Superfamily: L-lysine transport protein

Query Match 73.2%; Score 41; DB 2; Length 490;
 Best Local Similarity 50.0%; Pred. No. 78;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
 : ||| |||
 Db 92 WGYWLSAWAG 101

RESULT 9
 C95282
 A:Description: probable ABC transporter, periplasmic solute-binding protein, family 5 Sma0302 [imported]
 C:Species: Sinorhizobium meliloti
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 A:Accession: C95282
 R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bove
 ; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: C95282
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-508 <KUR>
 A:Cross-references: UNIPROT:Q930N2; GB:AE006469; PIDN:AAK64821.1; PID:g14523232; GSPDB:G
 A:Experimental source: strain 1021, megaplasmid pSymA
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebaullt, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Sma0302
 A:Genome: plasmid

Query Match 73.2%; Score 41; DB 2; Length 508;
 Best Local Similarity 75.0%; Pred. No. 80;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTW 8
 : ||| |||
 Db 392 EGYWDTIW 399

RESULT 10
 C75176
 A:Description: na+/h+ antiporter (napa-2) PAB0390 - Pyrococcus abyssi (strain Orsay)
 C:Species: Pyrococcus abyssi
 C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 A:Accession: C75176
 R:Anonymous, Genoscope
 submitted to the EMBL Data Library, July 1999
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc
 A:Reference number: A75001
 A:Accession: C75176
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-426 <NAW>
 A:Cross-references: UNIPROT:Q9V160; GB:AJ248284; GB:AL096836; NID:g5457730; PIDN:CAB4949
 A:Experimental source: strain Orsay
 C:Genetics:
 A:Gene: napa-2; PAB0390
 C:Superfamily: Aquifex aeolicus Na+/H+-exchanging protein napA1

Query Match 71.4%; Score 40; DB 2; Length 426;
 Best Local Similarity 50.0%; Pred. No. 96;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
 : ||| |||
 Db 319 FGAWITAWG 328

RESULT 11
 T46745
 A:Description: arginine/ornithine antiporter [imported] - Lactobacillus sakei
 C:Species: Lactobacillus sakei
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
 A:Accession: T46745
 R:Zuniga, M.; Champomier-Verges, M.C.; Perez-Martinez, G.; Zagorec, M.; Perez-Martinez,
 J. Bacteriol. 180, 4154-4159, 1998
 A:Title: Structural and functional analysis of the gene cluster encoding the enzymes of
 A:Reference number: Z23141; MUID:98361904; PMID:9696763
 A:Accession: T46745
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-475 <ZUN>
 A:Cross-references: UNIPROT:O53092; EMBL:AJ001330; NID:g2764610; PIDN:CAA04686.1; PID:g2
 C:Genetics:
 A:Gene: arcD
 C:Function:
 A:Description: necessary for arginine transport; involved in ornithine-arginine exchange
 A:Pathway: arginine catabolism
 C:Superfamily: L-lysine transport protein

Query Match 71.4%; Score 40; DB 2; Length 475;
 Best Local Similarity 62.5%; Pred. No. 1.1e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTW 8
 : ||| |||
 Db 93 WGYWLSAW 100

RESULT 12
 T50103
 A:Description: probable oxidoreductase [imported] - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
 A:Accession: T50103
 R:Seeger, K.; Harris, D.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, January 2000

A:Reference number: Z25038

A:Accession: T50103

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-583 <SEE>

A:Cross-references: UNIPROT:Q9US28; EMBL:AL136499; PIDN:CAB66164.1; GSPDB:GN000666; SPDB:

A:Experimental source: strain 972h(-); cosmid c1783

C:Genetics:

A:Gene: SPDB:SPAC1783.01

A:Map position: 1

Query Match 71.4%; Score 40; DB 2; Length 583;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:|:|:|:|:
Db 243 GGRWPTWGE 252

RESULT 13

Ins P4-binding protein - human

C:Species: Homo sapiens (man)

C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 04-Apr-2004

C:Accession: S58888

R:Cullen, P.J.; Hsuan, J.J.; Truong, O.; Letcher, A.J.; Jackson, T.R.; Dawson, A.P.; Irv

Nature 376, 527-530, 1995

A:Title: Identification of a specific Ins(1,3,4,5)P(4)-binding protein as a member of th

A:Reference number: S58888; MUID:95364929; PMID:7637787

A:Accession: S58888

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-829 <CUL>

A:Cross-references: EMBL:X89399

F:328-539/Domain: ras-specific GAP catalytic domain homology <GAP>

F:571-670/Domain: pleckstrin repeat homology <PLK>

Query Match 71.4%; Score 40; DB 2; Length 829;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:|:|:|:|:
Db 688 SGHWLCCWGA 697

RESULT 14

Ins P4-binding protein - human

C:Species: Homo sapiens (man)

C:Date: 05-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 04-Apr-2004

C:Accession: S71847

R:Cullen, P.J.

submitted to the EMBL Data Library, July 1995

A:Reference number: S71847

A:Accession: S71847

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-829 <CUL>

A:Cross-references: EMBL:X89399; NID:g963051; PID:g963052

F:328-539/Domain: ras-specific GAP catalytic domain homology <GAP>

F:571-670/Domain: pleckstrin repeat homology <PLK>

Query Match 71.4%; Score 40; DB 2; Length 829;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
:|:|:|:|:
Db 688 SGHWLCCWGA 697

RESULT 15

A33830

cation efflux system membrane protein czca - Alcaligenes eutrophus

C:Species: Alcaligenes eutrophus

C:Date: 23-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 09-Jul-2004

C:Accession: A33830

R:Nies, D.H.; Nies, A.; Chu, L.; Silver, S.

Proc. Natl. Acad. Sci. U.S.A. 86, 7351-7355, 1989

A:Title: Expression and nucleotide sequence of a plasmid-determined divalent cation efflu

A:Reference number: A33830; MUID:90017477; PMID:2678100

A:Accession: A33830

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1063 <NIE>

A:Cross-references: UNIPROT:P13511; GB:M26073

C:Superfamily: cation efflux system membrane protein czca

C:Keywords: transmembrane protein

Query Match 70.5%; Score 39.5; DB 2; Length 1063;
Best Local Similarity 60.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 1 XGYWLTWGX 10
:|:|:|:|:
Db 861 AGYWWT-WGG 869

Search completed: January 3, 2005, 16:34:51
Job time : 23.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:12:00 ; Search time 121.6 Seconds
(without alignments)
47.317 Million cell updates/sec

Title: US-10-046-922-34
Perfect score: 56
Sequence: 1 XGYWLTWGX 10

Scoring table: BLOSUM62DX 1
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	80.4	221	Q74JK6	Q74jk6 lactobacill
2	45	80.4	221	AAS08923	AAS08923 lactobaci
3	44	78.6	284	O6MC01	O6mc01 parachlamyd
4	44	78.6	284	CAF23898	CAF23898 parachlam
5	44	78.6	441	Q922T6	Q922t6 rhizobium m
6	44	78.6	733	Q8H2N7	Q8h2n7 oryza sativ
7	44	78.6	1498	MRP6_MOUSE	Q91s7 mus musculu
8	44	78.6	1502	MRP6_RAT	Q88269 rattus norv
9	44	78.6	1502	MRP6_HUMAN	Q95255 homo sapien
10	43	76.8	286	Q82285	Q82285 enterococcu
11	43	76.8	376	Q6CBE4	Q6cbe4 yarrowia li
12	43	76.8	413	Q9HKA9	Q9hka9 thermoplasm
13	43	76.8	419	Q97VB7	Q97vb7 sulfolobus
14	43	76.8	1052	Q7Y284	Q7y284 plasmodium
15	43	76.8	1308	Q8T6H2	Q8t6h2 dictyosteli
16	43	76.8	1652	Q74DR3	Q74dr3 geobacter s
17	43	76.8	1652	AAR34628	AAR34628 geobacter
18	42	75.0	401	Q9KHQ6	Q9khq6 bacteroides
19	42	75.0	425	O59255	O59255 pyrococcus
20	42	75.0	736	Q8DMV4	Q8dmv4 synecococc
21	42	75.0	1970	Q88H71	Q88h71 pseudomonas
22	41	73.2	117	Q7S013	Q7s013 neurospora
23	41	73.2	167	OB_MACMU	Q28504 macaca mula
24	41	73.2	253	Q32816	Q32816 lactococcus
25	41	73.2	312	Q6C129	Q6c129 yarrowia li
26	41	73.2	331	Q775C7	Q775c7 bordetella
27	41	73.2	331	AAR57683	AAR57683 bordeteli
28	41	73.2	376	Q8YX40	Q8yx40 anabaena sp
29	41	73.2	432	Q8ZUS8	Q8zus8 pyrobaculum
30	41	73.2	443	Q7NSJ2	Q7nsj2 chromobacte
31	41	73.2	459	Q9KGV3	Q9kgv3 lactococcus

32 41 73.2 482 2 Q89IX1
33 41 73.2 490 2 Q9CE15
34 41 73.2 508 2 Q93ON2
35 41 73.2 526 2 Q9K574
36 40 71.4 91 2 Q82A22
37 40 71.4 131 2 Q7QUC2
38 40 71.4 144 2 Q6M566
39 40 71.4 144 2 CAF21531
40 40 71.4 162 2 Q8NOB6
41 40 71.4 189 2 Q8EBD7
42 40 71.4 257 2 Q9ARZ0
43 40 71.4 294 1 RBSU_LACPL
44 40 71.4 426 2 Q9V160
45 40 71.4 427 2 Q8U0I6

ALIGNMENTS

RESULT 1
Q74JK6 PRELIMINARY; PRT; 221 AA.
AC Q74JK6;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hemolysin-like protein.
GN OrderedLocusNames-LJ1101;
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Fridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwahlen M.-C., Rouvet M., Altermann E., Barrangu R.,
RA Mollet B., Mercenier A., Klaenhammer T., Arigoni F., Schell M.A.;
RT "The genome sequence of the probiotic intestinal bacterium
RT Lactobacillus johnsonii NCC 533.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
DR EMBL; AB017203; AAS08923.1; -.
DR InterPro; IPR004254; HlyVIII-related.
DR InterPro; IPR005744; HlyIII-
DR Pfam; PF03006; HlyVIII; 1.
DR TIGRFAMS; TIGR01065; hlyVIII; 1.
KW Complete proteome.
SQ SEQUENCE 221 AA; 24721 NW; 1DBC78FF9810E152 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 221;
Best Local Similarity 50.0%; Pred. No. 54;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
Db 171 TGFLLVWGG 180

RESULT 2
AAS08923 PRELIMINARY; PRT; 221 AA.
AC AAS08923;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE Hemolysin-like protein.
GN LJ1101.
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]_

RP	SEQUENCE FROM N.A.
RC	STRAIN=NCC 533;
RX	PubMed=14966310;
RA	Pridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA	Pittet A.-C., Zwhlen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA	Mollet B., Mercenier A., Kiehnhammer T., Arigoni F., Schell M.A.;
RT	"The genome sequence of the probiotic intestinal bacterium
RT	Lactobacillus johnsonii NCC 533.";
RL	Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
DR	ENBL: AE017203; AAS08923.1;
SQ	SEQUENCE 221 AA; 24721 MW; IDBC78FF9810E152 CRC64;
Query Match	80.4%; Score 45; DB 2; Length 221;
Best Local Similarity	50.0%; Pred. No. 54;
Matches	5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Qy	1 XGYWLTWGKX 10
Dy	: : :
Db	171 TGFLLVGG 180
RESULT 3	
Q6MC01	PRELIMINARY; PRT; 284 AA.
ID	Q6MC01
AC	Q6MC01;
DT	05-JUL-2004 (TrEMBLrel. 27, Created)
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE	Hypothetical protein.
GN	ORFNames=pc1174;
OS	Parachlamydia sp. (strain UAE25) (subsp. Acanthamoeba sp.).
OC	Bacteria; Chlamydiae; Chlamydiales; Parachlamydiaceae; Parachlamydia.
OX	NCBI_TaxID=264201;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=null;
RA	Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhold U.,
RA	Fartmann B., Brandt P., Nyakatura G.J., Droge M., Frishman D.,
RA	Rattei T., Mewes H.-W., Wagner M.;
RT	"Genome sequence of an amoeba symbiont and its use for reconstructing
RT	the evolutionary history of chlamydiae.";
RL	Submitted (JAN-2003) to the ENBL/GenBank/DBJ databases.
DR	ENBL: BX908798; CAF23898.1;
KW	Hypothetical protein.
SQ	SEQUENCE 284 AA; 33507 MW; 569E29FD07F6AECA CRC64;
Query Match	78.6%; Score 44; DB 2; Length 284;
Best Local Similarity	50.0%; Pred. No. 97;
Matches	5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Qy	1 XGYWLTWGKX 10
Dy	: : :
Db	271 QAYWLSLWGI 280
RESULT 4	
CAF23898	PRELIMINARY; PRT; 284 AA.
ID	CAF23898
AC	CAF23898;
DT	14-APR-2004 (TrEMBLrel. 27, Created)
DT	14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT	14-APR-2004 (TrEMBLrel. 27, Last annotation update)
DE	Hypothetical protein.
GN	PC1174.
OS	Parachlamydia sp. UAE25.
OC	Bacteria; Chlamydiae; Chlamydiales; Parachlamydiaceae; Parachlamydia.
OX	NCBI_TaxID=264201;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=UWE25;
RA	Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhold U.,
RA	Fartmann B., Brandt P., Nyakatura G.J., Droge M., Frishman D.,
RA	Rattei T., Mewes H.-W., Wagner M.;


```

DR Pfam; PF00324; AA_permease; 1.
KW Complete proteome; Plasmod; Transmembrane; Transport.
SQ SEQUENCE 441 AA; 46277 MW; B6F5C9EC96011AD0 CRC64;

Query Match 78.6%; Score 44; DB 2; Length 441;
Best Local Similarity 50.0%; Pred. NO. 1.4e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYLTIWGX 10
DB 91 WGYWISIWAS 100

RESULT 6
Q8H2N7 PRELIMINARY; PRT; 733 AA.
AC Q8H2N7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein QJ1138_B05.118.
GN Name=QJ1138_B05.118;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005486; BAC16197.1;
DR Gramene; Q8H2N7;
DR InterPro; IPR007658; DUF594.
DR Pfam; PF04578; DUF594; 1.
KW Hypothetical protein.
SQ SEQUENCE 733 AA; 82939 MW; E95884DAD1DC2AC9 CRC64;

Query Match 78.6%; Score 44; DB 2; Length 733;
Best Local Similarity 75.0%; Pred. NO. 2.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 YWLTWGX 10
DB 328 YWTTWGR 335

RESULT 7
MRP6 MOUSE
ID MRP6 MOUSE STANDARD; PRT; 1498 AA.
AC Q9R1S7; Q8OVB6;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Multidrug resistance-associated protein 6 (ATP-binding cassette, sub-
DE family C, member 6).
GN Name=Abcc6; Synonyms=Mrp6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ddY; TISSUE=Liver;
RA Morikawa A., Suzuki H., Hirohashi T., Sugiyama Y.;
RT "Mus musculus mRNA for multidrug resistance-associated protein 6
RT (MRP6), complete cds.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 685-1498 FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP REVIEW.
RX DOI=10.1038/mp.a002623.01;
RA Beck C., Le Saux O., Varadi A., Boyd C.;
RA "Abcc6.";
RL (er) AFCS-Nature Molecule Pages (2004).
CC -!- FUNCTION: May participate directly in the active transport of
CC drugs into subcellular organelles or influence drug distribution
CC indirectly (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB028737; BAA83820.1; --
CC EMBL; BC049980; AAH49980.1; --
CC HSSP; Q03518; 1J07.
CC MGD; MGI:1351634; Abcc6.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001140; ABC_TM_transp.
CC InterPro; IPR003439; ABC_transporter.
CC InterPro; IPR005292; MRP_assoc.
CC Pfam; PF00664; ABC_membrane; 2.
CC Pfam; PF00005; ABC_tran; 2.
CC ProDom; PD000006; ABC_transporter; 1.
CC SMART; SM00382; AAA; 2.
CC TIGRfam; TIGR00957; MRP_assoc_pro; 1.
CC PROSITE; PS00929; ABC_TMIF; 2.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Glycoprotein; Repeat; Transmembrane; Transport.
FT DOMAIN 1 37 Extracellular (By similarity).
FT TRANSMEM 38 58 1 (By similarity).
FT DOMAIN 59 78 Cytoplasmic (By similarity).
FT TRANSMEM 79 99 2 (By similarity).
FT DOMAIN 100 104 Extracellular (By similarity).
FT TRANSMEM 105 125 3 (By similarity).
FT DOMAIN 126 137 Cytoplasmic (By similarity).
FT TRANSMEM 138 155 4 (By similarity).
FT DOMAIN 156 173 Extracellular (By similarity).
FT TRANSMEM 174 194 5 (By similarity).
FT DOMAIN 195 300 Cytoplasmic (By similarity).
FT TRANSMEM 301 321 6 (By similarity).
FT DOMAIN 322 347 Extracellular (By similarity).
FT TRANSMEM 348 368 7 (By similarity).
FT DOMAIN 369 424 Cytoplasmic (By similarity).
FT TRANSMEM 425 445 8 (By similarity).
FT DOMAIN 446 Extracellular (By similarity).

```

```

FT TRANSMEM 449 469 9 (By similarity).
FT DOMAIN 470 531 Cytoplasmic (By similarity).
FT TRANSMEM 532 552 10 (By similarity).
FT DOMAIN 553 574 Extracellular (By similarity).
FT TRANSMEM 575 595 11 (By similarity).
FT DOMAIN 596 934 Cytoplasmic (By similarity).
FT TRANSMEM 935 955 12 (By similarity).
FT DOMAIN 956 992 Extracellular (By similarity).
FT TRANSMEM 993 1013 13 (By similarity).
FT DOMAIN 1014 1056 Cytoplasmic (By similarity).
FT TRANSMEM 1057 1077 14 (By similarity).
FT DOMAIN 1078 1078 Extracellular (By similarity).
FT TRANSMEM 1079 1099 15 (By similarity).
FT DOMAIN 1100 1170 Cytoplasmic (By similarity).
FT TRANSMEM 1171 1191 16 (By similarity).
FT DOMAIN 1192 1193 Extracellular (By similarity).
FT TRANSMEM 1194 1214 17 (By similarity).
FT DOMAIN 1215 1498 Cytoplasmic (By similarity).
FT TRANSMEM 1499 1678 ABC transporter 1.
FT DOMAIN 1679 1944 ABC transporter 2.
FT NP_BIND 1945 1968 ATP (Potential).
FT NP_BIND 1969 1994 ATP (Potential).
FT CARBOHYD 1995 2021 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 2022 2048 N-linked (GlcNAc...) (Potential).
FT CONFLICT 2049 2066 A -> V (in Ref. 2).
FT CONFLICT 2067 2084 I -> T (in Ref. 2).
FT CONFLICT 2085 2102 H -> Q (in Ref. 2).
FT CONFLICT 2103 2120 L -> V (in Ref. 2).
FT CONFLICT 2121 2138 N -> S (in Ref. 2).
SQ SEQUENCE 1498 AA; 164788 MW; EFCFF33FOECB13C CRC64;

Query Match 78.6%; Score 44; DB 1; Length 1498;
Best Local Similarity 50.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLWIGX 10
Db 959 QGYWLSLWAD 968
:|||||:
:|||||:

RESULT 8
MRP6 RAT STANDARD; PRT; 1502 AA.
AC O88269;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Multidrug resistance-associated protein 6 (ATP-binding cassette, sub-
DE family C, member 6) (MRP-like protein-1) (MLP-1).
GN Name=Abcc6; Synonyms=Mrp6, Mlp1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RA MEDLINE=98279126; PubMed=9614210;
RA Hirohashi T., Suzuki H., Ito K., Ogawa K., Shimizu T.,
RA Sugiyama Y.;
RT "Hepatic expression of multidrug resistance-associated protein-like
RT proteins maintained in eels hyperbilirubinemic rats.";
RL Mol. Pharmacol. 53:1068-1075 (1998).
[2]
SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=20159081; PubMed=10692506;
RA Madon J., Hagenbuch B., Landmann L., Meier P.J., Stieger B.;
RT "Transport function and hepatocellular localization of mrp6 in rat
RT liver.";
RL Mol. Pharmacol. 57:634-641 (2000).
CC -!- FUNCTION: May participate directly in the active transport of
CC drugs into subcellular organelles or influence drug distribution

```

```

CC indirectly (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- TISSUE SPECIFICITY: High in liver and lower in duodenum and
CC kidney.
CC -!- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; AB010466; BAA28954.1; -.
CC EMBL; U73038; AAD12747.1; -.
CC F01; T42216; T42216.
CC HSSP; P08716; 1MT0.
CC RGD; 620268; Abcc6.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001140; ABC_TM_transp.
CC InterPro; IPR003439; ABC_transporter.
CC InterPro; IPR005232; MRP_assoc.
CC Pfam; PF00664; ABC_membrane; 2.
CC Pfam; PF00005; ABC_tran; 2.
CC ProDom; PD000006; ABC_transporter; 1.
CC SMART; SM00382; AAA; 2.
CC TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Glycoprotein; Repeat; Transmembrane; Transport.
FT DOMAIN 5 37 Extracellular (By similarity).
FT TRANSMEM 38 58 1 (By similarity).
FT DOMAIN 59 78 Cytoplasmic (By similarity).
FT TRANSMEM 79 99 2 (By similarity).
FT DOMAIN 100 104 Extracellular (By similarity).
FT TRANSMEM 105 125 3 (By similarity).
FT DOMAIN 126 137 Cytoplasmic (By similarity).
FT TRANSMEM 138 155 4 (By similarity).
FT DOMAIN 156 173 Extracellular (By similarity).
FT TRANSMEM 174 194 5 (By similarity).
FT DOMAIN 195 300 Cytoplasmic (By similarity).
FT TRANSMEM 301 321 6 (By similarity).
FT DOMAIN 322 347 Extracellular (By similarity).
FT TRANSMEM 348 368 7 (By similarity).
FT DOMAIN 369 424 Cytoplasmic (By similarity).
FT TRANSMEM 425 445 8 (By similarity).
FT DOMAIN 446 448 Extracellular (By similarity).
FT TRANSMEM 449 469 9 (By similarity).
FT DOMAIN 470 531 Cytoplasmic (By similarity).
FT TRANSMEM 532 552 10 (By similarity).
FT DOMAIN 553 574 Extracellular (By similarity).
FT TRANSMEM 575 595 11 (By similarity).
FT DOMAIN 596 938 Cytoplasmic (By similarity).
FT TRANSMEM 939 959 12 (By similarity).
FT DOMAIN 960 996 Extracellular (By similarity).
FT TRANSMEM 997 1017 13 (By similarity).
FT DOMAIN 1018 1060 Cytoplasmic (By similarity).
FT TRANSMEM 1061 1081 14 (By similarity).
FT DOMAIN 1082 1082 Extracellular (By similarity).
FT TRANSMEM 1083 1103 15 (By similarity).
FT DOMAIN 1104 1174 Cytoplasmic (By similarity).
FT TRANSMEM 1175 1195 16 (By similarity).
FT DOMAIN 1196 1197 Extracellular (By similarity).
FT TRANSMEM 1198 1218 17 (By similarity).
FT DOMAIN 1219 1502 Cytoplasmic (By similarity).
FT TRANSMEM 1503 1502 ABC transporter 1.
FT DOMAIN 1503 1502 ABC transporter 2.
FT NP_BIND 1503 1502 ATP (Potential).
FT NP_BIND 1503 1502 ATP (Potential).
FT CARBOHYD 1503 1502 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1502 AA; 164995 MW; 539901B674A74A28 CRC64;

```

Query Match 78.6%; Score 44; DB 1; Length 1502;
 Best Local Similarity 50.0%; Pred. No. 4.4e+02;
 Matches 5; Conservative 1; Indels 0; Gaps 0;

QY 1 XGYWLIWG X 10
 :|||||:
 Db 963 QGYWLSLWAD 972

RESULT 9
 MRP6_HUMAN STANDARD; PRT; 1503 AA.
 AC Q95255; P78420; Q9UMZ7;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last sequence update)
 DE Multidrug resistance-associated protein 6 (ATP-binding cassette, sub-
 family C, member 6) (Anthracycline resistance-associated protein)
 DE (Multi-specific organic anion transporter-E) (MOAT-E).
 GN Name=ABCC6; Synonyms=MRP6, ARA;
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99107222; PubMed=9892204;
 RA Kool M., van der Linden M., de Haas M., Baas P., Borst P.;
 RT "Expression of human MRP6, a homologue of the multidrug resistance
 RT protein gene MRP1, in tissues and cancer cells.";
 RL Cancer Res. 59:175-182(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99425270; PubMed=10493829;
 RA Loftus B.J., Mason U.-J., Shedd V.P., Kalush F., Brandon R.,
 RA Fuhrmann J.J., Klein T., Crosby M.L., Barnstead M., Cronin L.,
 RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
 RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
 RT "Genome duplications and other features in 12 Mb of DNA sequence from
 RT human chromosome 16p and 16q.";
 RL Genomics 60:295-308(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99352020; PubMed=10424734;
 RA Belinsky M.G., Kruh G.D.;
 RT "MOAT-E (ARA) is a full-length MRP/cMOAT subfamily transporter
 RT expressed in kidney and liver.";
 RL Br. J. Cancer 80:1342-1349(1999).
 RN [4]
 RP FUNCTION, AND CHARACTERIZATION OF VARIANTS PXE PHE-1298; ARG-1302 AND
 RP SER-1321.
 RX MEDLINE=21988190; PubMed=11880368; DOI=10.1074/jbc.M10918200;
 RA Ilias A., Urban Z., Seidl T.B., Le Saux O., Sinko E., Boyd C.D.,
 RA Sarkadi B., Varadi A.;
 RT "Loss of ATP-dependent transport activity in pseudoxanthoma elasticum-
 RT associated mutants of human ABCC6 (MRP6).";
 RL J. Biol. Chem. 277:16860-16867(2002).
 RN [5]
 RP REVIEW, AND VARIANT PXE PRO-455.
 RX MEDLINE=21321623; PubMed=11427982; DOI=10.1016/S1471-4914(00)01869-4;
 RA Uitto J., Pulkkinen L., Ringpfeil F.;
 RT "Molecular genetics of pseudoxanthoma elasticum: a metabolic disorder
 RT at the environment-genome interface?";
 RL Trends Mol. Med. 7:13-17(2001).
 RN [6]
 RP VARIANT GLN-1268.
 RX MEDLINE=20374463; PubMed=10913334; DOI=10.1006/bbrc.2000.3101;
 RA Germain D.P., Perdu J., Remones V., Jeunenaitre X.;
 RT "Homozygosity for the R1268Q mutation in MRP6, the pseudoxanthoma
 RT elasticum gene, is not disease-causing.";
 RL Biochem. Biophys. Res. Commun. 274:297-301(2000).
 RN [7]

RP VARIANT ARG-64.
 RX MEDLINE=20514578; PubMed=11058917;
 RX DOI=10.1010/1098-1004(200011)16:5<449::AID-HUMU24>3.0.CO;2-O;
 RA Germain D.P., Perdu J., Remones V., Manzoni K., Jeunenaitre X.;
 RT "Identification of two polymorphisms (c189G>C; c190T>C) in exon 2 of
 RT the human MRP6 gene (ABCC6) by screening of pseudoxanthoma elasticum
 RT patients: possible sequence correction?";
 RL Hum. Mutat. 16:449-449(2000).
 RN [8]
 RP VARIANT PXE CYS-1339, AND VARIANT GLN-632.
 RX MEDLINE=20408303; PubMed=10954200;
 RA Strub B., Cai L., Zaech S., Ji W., Chung J., Lumsden A., Stumm M.,
 RA Huber M., Schaen L., Kim C.-A., Goldsmith L.A., Viljoen D.,
 RA Figuera L.E., Fuchs W., Munier F., Ramesar R., Hohl D., Richards R.,
 RA Neldner K.H., Lindpaintner K.;
 RT "Mutations of the gene encoding the transmembrane transporter protein
 RT ABC-C6 cause pseudoxanthoma elasticum.";
 RL J. Mol. Med. 78:282-286(2000).
 RN [9]
 RP VARIANTS PXE PRO-1114; GLN-1138 AND TRP-1314, AND VARIANT ALA-614.
 RX MEDLINE=20296630; PubMed=10835642; DOI=10.1038/76102;
 RA Le Saux O., Urban Z., Tschuch C., Csiszar K., Bacchelli B.,
 RA Quaglini D., Pasquali-Ronchetti I., Pope F.M., Richards A., Terry S.,
 RA Bercovitch L., de Paeppe A., Boyd C.D.;
 RT "Mutations in a gene encoding an ABC transporter cause pseudoxanthoma
 RT elasticum.";
 RL Nat. Genet. 25:223-227(2000).
 RN [10]
 RP VARIANT PXE TRP-1138, AND VARIANT GLN-1268.
 RX MEDLINE=20283940; PubMed=10811882; DOI=10.1073/pnas.100041297;
 RA Ringpfeil F., Leubwohl M.G., Cristiano A.M., Uitto J.;
 RT "Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a
 RT transmembrane ATP-binding cassette (ABC) transporter.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:6001-6006(2000).
 RN [11]
 RP VARIANTS PXE LYS-411; GLN-518; SER-568; PRO-673; GLN-765; PRO-1114;
 RP TRP-1121; PRO-1138; GLN-1138; ASP-1203; PHE-1298; ILE-1301; ARG-1302;
 RP PRO-1303; GLN-1314; TRP-1314; SER-1321; CYS-1339; HIS-1347; ASN-1361
 RP AND THR-1424, AND VARIANTS ASP-61; ARG-207; GLY-265; GLU-281; VAL-319;
 RP LYS-497; ALA-614; GLN-632; HIS-953; CYS-1241 AND GLN-1268.
 RX MEDLINE=21426347; PubMed=11536079;
 RA Le Saux O., Beck K., Sachsinger C., Silvestri C., Treiber C.,
 RA Pasquali-Ronchetti I., Johnson E.W., De Paeppe A., Pope F.M.,
 RA Pasquali-Ronchetti I., Bercovitch L., Terry S., Boyd C.D.;
 RT "A spectrum of ABCC6 mutations is responsible for pseudoxanthoma
 RT elasticum.";
 RL Am. J. Hum. Genet. 69:749-764(2001).
 RN [12]
 RP VARIANTS PXE 60-ARG--TYR-62 DEL; ARG-364 AND ARG-1354, AND VARIANT
 RP GLY-265.
 RX MEDLINE=21558431; PubMed=11702217; DOI=10.1007/s004390100582;
 RA Pulkkinen L., Nakano A., Ringpfeil F., Uitto J.;
 RT "Identification of ABCC6 pseudogenes on human chromosome 16p:
 RT implications for mutation detection in pseudoxanthoma elasticum.";
 RL Hum. Genet. 109:356-365(2001).
 RN [13]
 RP VARIANTS ALA-614; GLN-632 AND GLN-1268.
 RX MEDLINE=21632106; PubMed=1176382;
 RA Wang J., Near S., Young K., Connolly P.W., Hegele R.A.;
 RT "ABCC6 gene polymorphism associated with variation in plasma
 RT lipoproteins.";
 RL J. Hum. Genet. 46:699-705(2001).
 CC -!- FUNCTION: May participate directly in the active transport of
 CC drugs into subcellular organelles or influence drug distribution
 CC indirectly. Transports glutathione conjugates as Leukotriene-c4
 CC (LTC4) and N-ethylmaleimide S-glutathione (NEM-GS).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- TISSUE SPECIFICITY: Expressed in kidney and liver. Very low
 CC expression in other tissues.
 CC -!- DISEASE: Defects in ABCC6 are a cause of autosomal dominant
 CC pseudoxanthoma elasticum (AD-PXE) [MIM:177850]. PXE is a disorder
 CC characterized by calcification of elastic fibers in skin, arteries
 CC and retina that results in dermal lesions with associated laxity

CC and loss of elasticity, arterial insufficiency and retinal
 CC hemorrhages leading to macular degeneration.
 CC -1- DISEASE: Defects in ABC6 are a cause of autosomal recessive
 CC pseudoxanthoma elasticum (AR-PXE) [MIM:264800].
 CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC -1- DATABASE: NAME=Mutations of the ABC6 gene;
 CC NOTE=Retina International's Scientific Newsletter;
 CC WWW="http://www.retina-international.com/sci-news/abcc6mut.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF076622; AAC79696.1; -. SEQ.
 DR EMBL; U91318; AAC15785.1; ALT_SEQ.
 DR EMBL; AF168791; AAD51293.1; -.
 DR HSSP; P08716; 1MT0.
 DR Genew; HGNC:57; ABCC6.
 DR MIM; 603234; -.
 DR MIM; 177850; -.
 DR MIM; 264800; -.
 DR GO; GO:0005624; C:membrane fraction; TAS.
 DR GO; GO:0005886; C:plasma membrane; TAS.
 DR GO; GO:0005524; F:ATP binding; TAS.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; TAS.
 DR GO; GO:0005215; F:transporter activity; TAS.
 DR GO; GO:0042493; P:response to drug; TAS.
 DR GO; GO:0006810; P:transport; TAS.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR001140; ABC_TM transpt.
 DR InterPro; IPR003439; ABC_transporter.
 DR InterPro; IPR005292; MRP_assoc.
 DR Pfam; PF00684; ABC_membrane; 2.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD000006; ABC_transporter; 1.
 DR SMART; SM00382; AAA; 2.
 DR TIGRFAMS; TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE; PS50929; ABC_TMIF; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
 KW ATP-binding; Disease mutation; Glycoprotein; Polymorphism; Repeat;
 KW Transmembrane; Transport; Vision.
 FT DOMAIN 1 31 Extracellular (By similarity).
 FT TRANSMEM 32 52 1 (By similarity).
 FT DOMAIN 53 72 Cytoplasmic (By similarity).
 FT TRANSMEM 73 93 2 (By similarity).
 FT DOMAIN 94 98 Extracellular (By similarity).
 FT TRANSMEM 99 119 3 (By similarity).
 FT DOMAIN 120 131 Cytoplasmic (By similarity).
 FT TRANSMEM 132 149 4 (By similarity).
 FT DOMAIN 150 167 Extracellular (By similarity).
 FT TRANSMEM 168 188 5 (By similarity).
 FT DOMAIN 189 302 Cytoplasmic (By similarity).
 FT TRANSMEM 303 323 6 (By similarity).
 FT DOMAIN 324 349 Extracellular (By similarity).
 FT TRANSMEM 350 370 7 (By similarity).
 FT DOMAIN 371 426 Cytoplasmic (By similarity).
 FT TRANSMEM 427 447 8 (By similarity).
 FT DOMAIN 448 450 Extracellular (By similarity).
 FT TRANSMEM 451 471 9 (By similarity).
 Query Match 78.6%; Score 44; DB 1; Length 1503;
 Best Local Similarity 50.0%; Pred. No. 4.4e+02;
 Matches 5; Conservative 1; Indels 0; Gaps 0;
 QY 1 XGYWLTWG 10
 :|:|:|:|:|:

Db 964 RGYWLSLWAD 973
 RESULT 10
 Q82285 PRELIMINARY; PRT; 266 AA.
 AC Q82285
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=EF31185;
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=V583 / ATCC 700802;
 RX MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
 RA Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
 RA Khouri H.M., Utterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,
 RA Fraser C.M.;
 RT "Role of mobile DNA in the evolution of vancomycin-resistant
 RT Enterococcus faecalis.";
 RL Science 299:2071-2074 (2003).
 DR EMBL; AF016957; AAO82859.1; -.
 DR TIGR; EF3185; -.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 266 AA; 28291 MW; CD6E72C4DF555A36 CRC64;
 Query Match 76.8%; Score 43; DB 2; Length 266;
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 XGYWLTWG 10
 :|:|:|:|:|:
 Db 202 AGTWTLWGQ 211
 RESULT 11
 Q6CBE4 PRELIMINARY; PRT; 376 AA.
 AC Q6CBE4
 DT 01-OCT-2004 (TrEMBLrel. 28, Created)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Similar to sp|P38071|Saccharomyces cerevisiae YBR026c.
 GN ORFNames=YALI0C19624g;
 OS Yarrowia lipolytica (Candida lipolytica).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OX NCBI_TaxID=4952;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;
 RG GENOLEVURES;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E., V.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykassen C.,
 RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Kerrast A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
 RA Pellenz S., Potier S., Richard G.P., Straub M.L., Suleau A.,
 RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof B., Wirth B.,
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,

```
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382129; CAG82338.1; -.
SQ SEQUENCE 376 AA; 41206 MW; E085FF7C32379DCB CRC64;

Query Match 76.8%; Score 43; DB 2; Length 376;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
Db 300 KGIWLTWAD 309
:|||||:

RESULT 12
Q9HKA9 PRELIMINARY; PRT; 413 AA.
AC Q9HKA9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transport protein related protein.
GN OrderedLocusNames=ra0692;
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
[1]

SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513(2000).
DR EMBL; AL445065; CAC11830.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS50850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 413 AA; 45192 MW; 307093BC358063D4 CRC64;

Query Match 76.8%; Score 43; DB 2; Length 413;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
Db 353 IGFWETLWGI 362
:|||||:

RESULT 13
Q97VB7 PRELIMINARY; PRT; 419 AA.
AC Q97VB7;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Permease.
GN OrderedLocusNames=SSO2718;
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
```

```
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awayez M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Brauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AB006865; AAK42828.1; -.
DR PIR; E90446; E90446.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS50850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 419 AA; 46499 MW; 6DB6AB6B5C6DA267 CRC64;

Query Match 76.8%; Score 43; DB 2; Length 419;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
Db 364 IGFWETLWGI 373
:|||||:

RESULT 14
Q7YZ84 PRELIMINARY; PRT; 1052 AA.
AC Q7YZ84;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PC10107C.
GN ORFNames=PC10107c;
OS Plasmodium chabaudi chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31271;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AS;
RX MEDLINE=22672904; PubMed=12787350;
RA Fischer K., Chavchich M., Huestis R., Wilson D.W., Kemp D.J., Saul A.;
RT "Ten families of variant genes encoded in subtelomeric regions of
RT multiple chromosomes of Plasmodium chabaudi, a malaria species that
RT undergoes antigenic variation in the laboratory mouse.";
RL Mol. Microbiol. 48:1209-1223(2003).
DR EMBL; AY149028; AAO06132.1; -.
SQ SEQUENCE 1052 AA; 110704 MW; D6C5B3B3247C23B8 CRC64;

Query Match 76.8%; Score 43; DB 2; Length 1052;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10
Db 656 SGVSNLWGT 665
:|||||:

RESULT 15
Q8T6H2 PRELIMINARY; PRT; 1308 AA.
AC Q8T6H2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
```

```

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ABC transporter ABCC.7.
GN Name=abcC7;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ax4;
RA Anjard C., Loomis W.F.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF474339; AAL85710.1; -.
DR HSP; P08716; IWO.
DR DictyBase; DDB0201630; abcC7.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transpt.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS0929; ABC_TM1P; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS0893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 1308 AA; 149576 MW; 336120AC8F737AC1 CRC64;

Query Match 76.8%; Score 43; DB 2; Length 1308;
Best Local Similarity 75.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 YWLTWGX 10
Db 764 YWLTWSD 771

```

Search completed: January 3, 2005, 16:32:30
 Job time : 122.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:33:55 ; Search time 147 Seconds
(without alignments)
24.403 Million cell updates/sec

Title: us-10-046-922-34
Perfect score: 56
Sequence: 1 XGYWLTWGX 10

Scoring table: BLOSUM62DX1
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 444336

Minimum DB seq length: 0

Maximum-DB-seq-length: 10

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_23Sep04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	56	100.0	10	5 ABP53931	Abp53931 VEGFR-3 b
2	56	100.0	10	5 ABP53932	Abp53932 VEGFR-3 b
3	32.5	58.0	10	2 AAR14709	Aar14709 Labaditin
4	32	57.1	7	3 AAY76794	Aay76794 Somatost
5	32	57.1	7	5 ABP53418	Abp53418 Backbone
6	32	57.1	9	5 ABP53933	Abp53933 VEGFR-3 b
7	32	57.1	10	4 AAG95260	Aag95260 Human com
8	31	55.4	8	2 AAW97529	Aaw97529 Antigenic
9	31	55.4	9	4 AAU02369	Aau02369 HLA bindi
10	31	55.4	9	4 AAU02282	Aau02282 HLA bindi
11	31	55.4	9	8 ADN64476	Adn64476 HLA bindi
12	31	55.4	10	4 AAB99759	Aab99759 Rhesus D
13	31	55.4	10	4 AAB99758	Aab99758 Rhesus D
14	30	53.6	8	4 ABP22594	Abp22594 HIV A11 m
15	30	53.6	8	4 ABP14287	Abp14287 HIV A03 s
16	30	53.6	8	4 ABP18686	Abp18686 HIV B62 s
17	30	53.6	8	4 ABP20292	Abp20292 HIV A03 m
18	30	53.6	8	4 ABP19992	Abp19992 HIV A03 m
19	30	53.6	8	4 ABP12050	Abp12050 HIV A02 s
20	30	53.6	8	4 ABP14288	Abp14288 HIV A03 s
21	30	53.6	8	4 ABP16999	Abp16999 HIV B27 s
22	30	53.6	8	4 ABP22402	Abp22402 HIV A11 m
23	30	53.6	8	4 ADH62029	Adh62029 Cyclic an
24	30	53.6	8	8 ADP67844	Adp67844 Anti-micr
25	30	53.6	8	8 ADQ28067	Adq28067 Excluded

ALIGNMENTS

RESULT 1

ABP53931
ID ABP53931 standard; peptide; 10 AA.

XX AC ABP53931;

DT DT 09-JAN-2003 (first entry)

DE DE VEGFR-3 binding peptide SEQ ID NO:34.

XX VEGFR-3 binding peptide SEQ ID NO:34.
XX Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
KW cytotatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
KW vulnary; cell surface receptor; cancer; neovascularisation;
KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
XX diabetes; PDGF; platelet derived growth factor.
XX Homo sapiens.
OS Synthetic.

XX FH Key Location/Qualifiers
FT FT Misc-difference 1 /note= "any amino acid"
FT FT Misc-difference 10 /note= "any amino acid"

XX WO200257299-A2.

XX 25-JUL-2002.

XX 16-JAN-2002; 2002WO-1B000099.

XX 17-JAN-2001; 2001US-0262476P.

XX (LUDW-) LUDWIG INST CANCER RES.
(LICN) LICENTIA LTD.

XX PI Alitalo K, Koivunen E, Kubo H;

XX WPI; 2002-691521/74.

XX DR New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
such as cancer and diseases of neovascularization.

XX Claim 12; Page 80; 149pp; English.

XX The present invention describes an isolated peptide (I) that binds to and

Aar37115 HIV envel
Aar37113 HIV envel
Aar73059 Antigen f
Aaw43338 Immunogen
Aay47445 Immunogen
Aay66203 HLA-A2-b1
Aay73202 HIV-deriv
Aam22296 HIV pepti
Aam22420 HIV pepti
Aam23186 HIV pepti
Aam22908 HIV pepti
Aam22295 HIV pepti
Aam22907 HIV pepti
Aam22952 HIV pepti
Aam22953 HIV pepti
Abp25039 HIV A2 su
Abp25292 HIV CTL e
Abp17781 HIV B58 s
Abp12202 HIV A02 s
Abp15275 HIV A24 s

CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnery activities, and can be used in gene therapy.
 CC Compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
 CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 XX
 SQ Sequence 10 AA;
 Query Match 100.0%; Score 56; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.077;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XGYWLTWGX 10
 Db 1 XGYWLTWGX 10
 RESULT 2
 ID ABP53932 standard; peptide; 10 AA.
 AC ABP53932;
 XX
 XX 09-JAN-2003 (first entry)
 DT
 DE VEGFR-3 binding peptide SEQ ID NO:35.
 XX
 XX Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
 KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
 KW cytostatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
 KW vulnery; cell surface receptor; cancer; neovascularisation;
 KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
 KW diabetes; PDGF; platelet derived growth factor.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX WO200257299-A2.
 XX
 XX 25-JUL-2002.
 XX
 XX 16-JAN-2002; 2002WO-1800099.
 XX
 XX 17-JAN-2001; 2001US-0262476P.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA (LICN) LICENTIA LTD.
 XX
 XX Alitalo K, Koivunen E, Kubo H;
 XX WPI; 2002-691521/74.
 DR
 XX
 XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
 PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
 PT such as cancer and diseases of neovascularization.
 XX
 XX Claim 13; Page 80; 149pp; English.
 XX
 CC The present invention describes an isolated peptide (I) that binds to and
 CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnery activities, and can be used in gene therapy.
 CC Compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 XX
 SQ Sequence 10 AA;
 Query Match 100.0%; Score 56; DB 5; Length 10;
 Best Local Similarity 80.0%; Pred. No. 0.077;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 XGYWLTWGX 10
 Db 1 CGYWLTMGC 10
 RESULT 3
 ID AAR14709 standard; peptide; 10 AA.
 AC AAR14709;
 XX
 XX 25-MAR-2003 (revised)
 DT 24-JAN-1992 (first entry)
 XX
 DE Labaditin.
 XX
 KW Cyclic; immunoglobulin; autoimmune disease.
 XX
 OS Jatropa multifida.
 XX
 PN WO9116345-A.
 XX
 PD 31-OCT-1991.
 XX
 PF 23-APR-1990; 90US-00512796.
 XX
 PR 23-APR-1990; 90US-00512796.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 XX
 XX Labadie RP;
 PI
 XX WPI; 1991-339755/46.
 DR
 XX
 XX New cyclic peptide(s) e.g. labaditin, having IGG binding properties -
 PT used for IGG binding in mammals including humans and for treating
 PT inflammatory diseases including rheumatic and autoimmune diseases.
 XX
 XX Claim 11; Page 12; 15pp; English.
 XX
 XX The peptide is isolated from plants of the Euphorbiaceae family (see FEBS
 CC Letters 256, 1-2, 1989). It selectively inhibits the classical activation
 CC pathway of the complement system by binding to IGG. It can be used to
 CC treat autoimmune disorders e.g. rheu- matic and other inflammatory
 CC disorders, and immuno-complex-related diseases, e.g. extrinsic allergic
 CC alveolitis. It may also be used to isolate IGG from blood plasma or serum
 CC to reduce the level of IGG, or as immunoassay reagent to determine IGG
 CC levels in body fluids. See also AAR14710. (Updated on 25-MAR-2003 to
 CC correct PA field.)
 XX
 SQ Sequence 10 AA;
 Query Match 58.0%; Score 32.5; DB 2; Length 10;
 Best Local Similarity 50.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 1; Gaps 1;
 QY 1 XGYWLTWGX 10
 Db 1 AGYV-TWGT 9

Best Local Similarity 57.1%; Pred. No. 1.7e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

RESULT 4
ID AAY76794
XX AAY76794 standard; peptide; 7 AA.
AC AAY76794;
XX 20-APR-2000 (first entry)
XX Somatostatin analogue peptide 3181.
DE Somatostatin analogue; therapy; cyclic peptide; autoimmune disease;
XX endocrine disorder; cancer; diabetic-associated complication; diagnosis;
KW gastrointestinal disorder; inflammation; pancreatitis;
KW atherosclerosis; restenosis; post-surgical pain; VIP secretion inhibitor;
KW hormone-secreting tumour; hormone-dependent tumour; diarrhoea;
KW vasoactive intestinal peptide; non-insulin dependent diabetes mellitus.
XX Synthetic.
XX Key Location/Qualifiers
FH Key Location/Qualifiers
FT Misc-difference 3 /note= "D-form residue"
FT Modified-site 7 /note= "Trp-NH2"
FT Modified-site 7 /note= "Trp-NH2"
XX WO965508-A1.
PN 23-DEC-1999.
XX 15-JUN-1999; 99WO-IL000329.
XX 19-JUN-1998; 98US-00100360.
PR 02-DEC-1998; 98US-00203389.
XX (PEPT-) PEPTOR LTD.
XX Hornik V, Afargan MM, Gellerman G;
XX WPI; 2000-136988/12.
XX Cyclized somatostatin analogs for inhibiting growth hormone secretion
PT from anterior pituitary and as antiproliferative agents for the treatment
PT of tumors.
XX Example 11; Page 61; 82pp; English.
XX This sequence represents a somatostatin analogue of the invention. The
CC invention relates to a backbone cyclised somatostatin analogue that has
CC one building unit containing a nitrogen atom of the peptide backbone
CC connected to a bridging group comprising an amide, thioether, thioester
CC or disulphide. At least one building unit is connected via a bridging
CC group to form a cyclic structure with a moiety selected from a second
CC building unit, side chain of or N-terminal amino acid residue. A
CC composition containing the analogue may be used for preventing disorders
CC such as cancers, autoimmune diseases, endocrine disorders, diabetic-
CC associated complications, gastrointestinal disorders, inflammatory
CC diseases, pancreatitis, atherosclerosis, restenosis and post-surgical
CC pain. It may also be used for diagnosing cancer. The backbone cyclic
CC analogue is used for imaging the existence of metastases. Somatostatin
CC analogues can be used for the treatment of patients with hormone-secreting
CC and hormone-dependent tumours. They reduce diarrhoea through the
CC inhibition of vasoactive intestinal peptide (VIP) secretion and by direct
CC effect on intestinal secretion. Somatostatin analogues selective to type
CC 2 and 5 receptors may be used for treatment of non-insulin dependent
CC diabetes mellitus. They are useful for the prevention of atherosclerosis
CC and restenosis. The analogues are metabolically stable, selective in
CC their in-vivo activities and safe
XX Sequence 7 AA;
SQ Query Match 57.1%; Score 32; DB 3; Length 7;

QY 2 GYWLTIW 8
Db 1 GYWKVCW 7
RESULT 5
ABP53418
ID ABP53418 standard; peptide; 7 AA.
XX ABP53418;
AC ABP53418;
XX 19-NOV-2002 (first entry)
XX Backbone cyclised somatostatin analogue PTR 3181.
XX Backbone cyclised somatostatin analogue; somatostatin; SRIF; analgesic;
KW somatotropin release inhibiting factor; somatostatin receptor subtype;
KW synthesis; antiarteriosclerotic; immunosuppressive; cytostatic; cancer;
KW antidiabetic; antiinflammatory; somatostatin receptor ligand;
KW atherosclerosis; autoimmune disease; diabetic-associated complication;
KW endocrine disorder; inflammation; gastrointestinal disorder; restenosis;
KW pancreatitis; post-surgical pain.
XX Synthetic.
XX Key Location/Qualifiers
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminally modified with Fmoc
(fluorenylmethoxycarbonyl)"
FT Misc-difference 3 /note= "D form residue"
FT Modified-site 7 /note= "amidated"
XX US2002052315-A1.
XX 02-MAY-2002.
XX 13-DEC-2000; 2000US-00734583.
XX 19-JUN-1998; 98US-00100360.
PR 02-DEC-1998; 98US-00203389.
PR 15-JUN-1999; 99WO-IL000329.
XX (HORN/) HORN V.
PA (AFAR/) AFARGAN M M.
PA (GELL/) GELLERMAN G.
XX Hornik V, Afargan MM, Gellerman G;
XX WPI; 2002-681319/73.
XX New backbone cyclized somatostatin analogs are e.g. useful in the
PT treatment of atherosclerosis, autoimmune diseases and cancers.
XX Example 12; Page 21; 30pp; English.
XX The present invention describes backbone cyclised somatostatin analogues
CC (I) that incorporates at least one building unit containing one nitrogen
CC atom of the peptide backbone connected to a bridging group (comprising an
CC amide, thioether, thioester or disulfide) where at least one building
CC unit is connected via the bridging group to form a cyclic structure with
CC a moiety selected from the group consisting of a second building unit,
CC the side chain of an amino acid residue of the sequence or the N-terminal
CC amino acid residue. (I) has antiarteriosclerotic, immunosuppressive,
CC cytostatic, antidiabetic, antiinflammatory and analgesic activities, and
CC can be used as a somatostatin receptor ligand. (I) are useful in the
CC treatment of atherosclerosis, autoimmune diseases, cancers, diabetic-
CC associated complications, endocrine disorders, inflammation,
CC gastrointestinal disorders, pancreatitis, post-surgical pain, and

CC restenosis. (I) can also be used in the diagnosis of cancer, by imaging
CC the existence of metastases, it being labeled with a detectable probe.
CC The present sequence represents a backbone cyclised somatostatin analogue
CC from the present invention

XX SQ Sequence 7 AA;

Query Match 57.1%; Score 32; DB 5; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.7e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GYWLTIW 8
| | | | |
Db 1 GYWKVCW 7

RESULT 6
ABP53933
ID ABP53933 standard; peptide; 9 AA.

XX AC ABP53933;

XX DT 09-JAN-2003 (first entry)

XX VEGFR-3 binding peptide SEQ ID NO:36.

XX Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
KW cytostatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
KW vulnary; cell surface receptor; cancer; neovascularisation;
KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
KW diabetes; PDGF; platelet derived growth factor.

XX OS Homo sapiens.
OS Synthetic.

XX PN WO200257299-A2.

XX PD 25-JUL-2002.

XX PF 16-JAN-2002; 2002WO-IB000099.

XX PR 17-JAN-2001; 2001US-0262476P.

XX PA (LUDW-) LUDWIG INST CANCER RES.
PA (LICN) LICENTIA LTD.

XX PI Alitalo K, Koivunen E, Kubo H;

XX WI; 2002-691521/74.

XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
PT such as cancer and diseases of neovascularization.

XX Claim 14; Page 80; 149pp; English.

XX The present invention describes an isolated peptide (I) that binds to and
CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
CC antidiabetic and vulnary activities, and can be used in gene therapy.
CC Compositions and methods from the present invention are useful for
CC diagnosing, evaluating and treating disorders mediated by the activity of
CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
CC liver, spleen, kidney, lymph node, small intestine, blood cells,
CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
CC chronic hepatitis, haemangiomas and diabetes. The present sequence
CC represents a specifically claimed VEGFR-3 binding peptide from the
CC present invention

XX SQ Sequence 9 AA;

Query Match 57.1%; Score 32; DB 5; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.7e+06;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 XGYWLTIW 8
| | | | |
Db 1 SGYVWDTW 8

RESULT 7
AAG95260
ID AAG95260 standard; peptide; 10 AA.
XX AC AAG95260;
XX DT 18-SEP-2001 (first entry)
XX Human complementary peptide, SEQ ID NO: 1454.
XX Human; complementary peptide; ligand; drug discovery; drug design.

XX OS Homo sapiens.

XX PN WO200142277-A2.

XX PD 14-JUN-2001.

XX PF 13-DEC-2000; 2000WO-GB004776.

XX PR 13-DEC-1999; 99GB-00029464.

XX PA (PROT-) PROTEOM LTD.

XX PI Roberts GW, Heal JR;

XX WI; 2001-408419/43.

XX A set of peptide ligands consisting of specific complementary peptides to
PT proteins encoded by genes of the human genome, useful in an assay for
PT screening and identifying of one or more novel peptides which are drug
PT candidates or pro-drugs.

XX Example 4; Page 250; 646pp; English.

XX The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification

XX SQ Sequence 10 AA;

Query Match 57.1%; Score 32; DB 4; Length 10;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYWLTIWGX 10
| | | | |
Db 1 GYWLTVIGG 9

RESULT 8
AAW97529
ID AAW97529 standard; peptide; 8 AA.

XX AC AAW97529;

XX DT 19-MAY-1999 (first entry)

XX Antigenic site of HN protein loop beta-4L23.

KW Antigenic site; haemagglutinin-neuraminidase; HN; paramyxoviridae;
 KW virus epitope; attachment protein; vaccine; immunodominant epitope.
 XX Measles virus.
 OS
 XX
 PN WO9902695-A2.
 XX
 PD 21-JAN-1999.
 XX
 XX
 PF 08-JUL-1998; 98WO-NL000390.
 XX
 XX
 PR 08-JUL-1997; 97EP-00202100.
 XX
 XX (DIER-) STICHTING INST DIERHOUDERIJ EN DIERGEZON.
 XX
 XX Langedijk JPM, Van Oirschot JT;
 XX WPI; 1999-120896/10.
 DR
 XX
 XX Isolated proteinaceous substance - comprising at least one virus epitope
 PT derived from an attachment protein of a paramyxovirus.
 PT
 PS Disclosure; Page 48; 63pp; English.
 XX
 XX AAW97452-571 represent antigenic sites derived from the haemagglutinin-
 CC neuraminidase (HN) protein of the paramyxoviridae. The specification
 CC describes 3-D models identifying a proteinaceous substance comprising at
 CC least one virus epitope derived from the attachment protein, which
 CC corresponds to an antigenic site present on one of the loops of HN. The
 CC antigenic sites can be used to produce vaccines, to detect the viruses,
 CC and to select the immunodominant epitope
 XX
 XX Sequence 8 AA;
 SQ

Query Match 55.4%; Score 31; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 YWLTII 7
 |||||
 Db 4 YWLTII 8

RESULT 9
 AAU02369
 ID AAU02369 standard; peptide; 9 AA.
 XX
 AC AAU02369;
 XX
 XX 29-AUG-2001 (first entry)
 DT
 XX
 DE HLA binding TADG-16 peptide #145.
 XX
 KW Human; extracellular serine protease; tumour antigen derived gene-16;
 KW TADG-16; ovarian carcinoma; breast cancer; lung cancer; colon cancer;
 KW prostate cancer; HLA type.
 XX
 XX Homo sapiens.
 OS
 XX WO200127257-A1.
 PN
 XX 19-APR-2001.
 PD
 XX 13-OCT-2000; 2000WO-US028558.
 PF
 XX 14-OCT-1999; 99US-00418527.
 PR
 XX (UYAR-) UNIV ARKANSAS.
 PA
 XX O'Brien TJ, Underwood LJ, Shigemasa K;
 PI WPI; 2001-273769/28.
 DR
 XX
 XX New tumor antigen-derived gene-16 protein, useful for diagnosis and
 PT treatment of ovarian, breast, lung, colon and prostate cancer.
 XX
 XX Example 8; Page 53; 124pp; English.
 PS
 XX AAU02225-AAU02384 represent TADG-16 peptides which are tested for their
 CC binding affinity to the 8 haplotypes HLA A0201, HLA A0205, HLA A1, HLA
 CC A24, HLA B7, HLA B8, HLA B2702, and HLA B403. Tumour antigen derived
 CC gene-16 protein, TADG-16 (AAU02223), is a novel human extracellular
 CC serine protease. TADG-16 is expressed in normal ovaries and testes and in
 CC certain ovarian carcinomas. TADG-16 contains the conserved catalytic
 CC triad, His-Asp-Ser, and a signal secretion sequence characteristic of the
 CC serine protease family. An antisense oligonucleotide having a
 CC complementary sequence to the TADG-16 nucleic acid is useful for treating
 CC various cancers, including ovarian, breast, lung, colon and prostate. The
 CC TADG-16 nucleic acid, TADG-16 protein and antibodies specific to TADG-16
 CC are useful for the diagnosis of cancer. TADG-16 protein or its fragments
 CC are useful for vaccinating an individual against TADG-16
 XX
 XX Sequence 9 AA;
 SQ

Query Match 55.4%; Score 31; DB 4; Length 9;
 Best Local Similarity 57.1%; Pred. No. 1.7e+06;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 WLTWGX 10
 |||||
 Db 3 WWTGWGY 9

RESULT 10
 AAU02282
 ID AAU02282 standard; peptide; 9 AA.
 XX
 AC AAU02282;
 XX
 XX 29-AUG-2001 (first entry)
 DT
 XX
 DE HLA binding TADG-16 peptide #58.
 XX
 KW Human; extracellular serine protease; tumour antigen derived gene-16;
 KW TADG-16; ovarian carcinoma; breast cancer; lung cancer; colon cancer;
 KW prostate cancer; HLA type.
 XX
 XX Homo sapiens.
 OS
 XX WO200127257-A1.
 PN
 XX 19-APR-2001.
 PD
 XX 13-OCT-2000; 2000WO-US028558.
 PF
 XX 14-OCT-1999; 99US-00418527.
 PR
 XX (UYAR-) UNIV ARKANSAS.
 PA
 XX O'Brien TJ, Underwood LJ, Shigemasa K;
 PI WPI; 2001-273769/28.
 DR
 XX
 XX New tumor antigen-derived gene-16 protein, useful for diagnosis and
 PT treatment of ovarian, breast, lung, colon and prostate cancer.
 XX
 XX Example 8; Page 53; 124pp; English.
 PS
 XX AAU02225-AAU02384 represent TADG-16 peptides which are tested for their
 CC binding affinity to the 8 haplotypes HLA A0201, HLA A0205, HLA A1, HLA
 CC A24, HLA B7, HLA B8, HLA B2702, and HLA B403. Tumour antigen derived
 CC gene-16 protein, TADG-16 (AAU02223), is a novel human extracellular
 CC serine protease. TADG-16 is expressed in normal ovaries and testes and in
 CC certain ovarian carcinomas. TADG-16 contains the conserved catalytic
 CC triad, His-Asp-Ser, and a signal secretion sequence characteristic of the
 CC serine protease family. An antisense oligonucleotide having a

PT New tumor antigen-derived gene-16 protein, useful for diagnosis and
 PT treatment of ovarian, breast, lung, colon and prostate cancer.
 XX
 PS Example 8; Page 55; 124pp; English.
 XX
 CC AAU02225-AAU02384 represent TADG-16 peptides which are tested for their
 CC binding affinity to the 8 haplotypes HLA A0201, HLA A0205, HLA A1, HLA
 CC A24, HLA B7, HLA B8, HLA B2702, and HLA B403. Tumour antigen derived
 CC gene-16 protein, TADG-16 (AAU02223), is a novel human extracellular
 CC serine protease. TADG-16 is expressed in normal ovaries and testes and in
 CC certain ovarian carcinomas. TADG-16 contains the conserved catalytic
 CC triad, His-Asp-Ser, and a signal secretion sequence characteristic of the
 CC serine protease family. An antisense oligonucleotide having a
 CC complementary sequence to the TADG-16 nucleic acid is useful for treating
 CC various cancers, including ovarian, breast, lung, colon and prostate. The
 CC TADG-16 nucleic acid, TADG-16 protein and antibodies specific to TADG-16
 CC are useful for the diagnosis of cancer. TADG-16 protein or its fragments
 CC are useful for vaccinating an individual against TADG-16
 XX
 XX Sequence 9 AA;
 SQ

Query Match 55.4%; Score 31; DB 4; Length 9;
 Best Local Similarity 57.1%; Pred. No. 1.7e+06;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 WLTWGX 10
 |||||
 Db 3 WWTGWGY 9

RESULT 10
 AAU02282
 ID AAU02282 standard; peptide; 9 AA.
 XX
 AC AAU02282;
 XX
 XX 29-AUG-2001 (first entry)
 DT
 XX
 DE HLA binding TADG-16 peptide #58.
 XX
 KW Human; extracellular serine protease; tumour antigen derived gene-16;
 KW TADG-16; ovarian carcinoma; breast cancer; lung cancer; colon cancer;
 KW prostate cancer; HLA type.
 XX
 XX Homo sapiens.
 OS
 XX WO200127257-A1.
 PN
 XX 19-APR-2001.
 PD
 XX 13-OCT-2000; 2000WO-US028558.
 PF
 XX 14-OCT-1999; 99US-00418527.
 PR
 XX (UYAR-) UNIV ARKANSAS.
 PA
 XX O'Brien TJ, Underwood LJ, Shigemasa K;
 PI WPI; 2001-273769/28.
 DR
 XX
 XX New tumor antigen-derived gene-16 protein, useful for diagnosis and
 PT treatment of ovarian, breast, lung, colon and prostate cancer.
 XX
 XX Example 8; Page 53; 124pp; English.
 PS
 XX AAU02225-AAU02384 represent TADG-16 peptides which are tested for their
 CC binding affinity to the 8 haplotypes HLA A0201, HLA A0205, HLA A1, HLA
 CC A24, HLA B7, HLA B8, HLA B2702, and HLA B403. Tumour antigen derived
 CC gene-16 protein, TADG-16 (AAU02223), is a novel human extracellular
 CC serine protease. TADG-16 is expressed in normal ovaries and testes and in
 CC certain ovarian carcinomas. TADG-16 contains the conserved catalytic
 CC triad, His-Asp-Ser, and a signal secretion sequence characteristic of the
 CC serine protease family. An antisense oligonucleotide having a

CC complementary sequence to the TADG-16 nucleic acid is useful for treating
 CC various cancers, including ovarian, breast, lung, colon and prostate. The
 CC TADG-16 nucleic acid, TADG-16 protein and antibodies specific to TADG-16
 CC are useful for the diagnosis of cancer. TADG-16 protein or its fragments
 CC are useful for vaccinating an individual against TADG-16

XX Sequence 9 AA;
 SQ

Query Match 55.4%; Score 31; DB 4; Length 9;
 Best Local Similarity 57.1%; Pred. No. 1.7e+06;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 WLTWGX 10
 | : | | :
 Db 1 WVTGWGY 7

RESULT 11
 ADN64476
 ID ADN64476 standard; peptide; 9 AA.

XX AC ADN64476;

XX DT 01-JUL-2004 (first entry)

XX DE HLA binding peptide #1076.

XX cytotactic; hepatotropic; virucide; antiinflammatory; anti-HIV;
 KW gene therapy; vaccine; HLA binding peptide; HTL epitope; liposome;
 KW prostate specific antigen; prostate specific membrane antigen;
 KW hepatitis B virus antigen; hepatitis C virus antigen;
 KW malignant melanoma antigen; MAGE; Epstein Barr virus; cancer;
 KW prostate cancer; AIDS; renal carcinoma; cervical carcinoma; lymphoma;
 KW chondyroma acuminatum.

XX Unidentified.

XX PN W02004031211-A2.

XX PD 15-APR-2004.

XX PF 03-OCT-2003; 2003WO-US031308.

XX PR 03-OCT-2002; 2002US-0416207P.

XX PR 08-OCT-2002; 2002US-0417269P.

XX PA (EPIM-) EPIMUNE INC.

XX PI Sidney J, Southwood S, Sette A;

XX WPI; 2004-347953/32.

XX New composition of peptides and nucleic acids capable of binding Major
 CC Histocompatibility Complex molecules, useful for diagnosing, preventing
 CC or treating viral infections or cancer, such as prostate cancer,
 CC hepatitis B or AIDS.

XX Claim 1; SEQ ID NO 1076; 186pp; English.

XX The invention relates to a novel composition comprising one or more
 CC peptides or nucleic acids encoding an HLA binding peptide. The
 CC composition further comprises an HTL epitope. It also comprises a spacer
 CC molecule, a carrier, an MHC targeting sequence or a lipid. The peptides
 CC are incorporated as part of a liposome. The peptide is from an antigen
 CC selected from prostate specific antigen (PSA), prostate specific membrane
 CC antigen (PSM), hepatitis B virus (HBV) antigen, hepatitis C virus (HCV)
 CC antigen, malignant melanoma antigen (MAGE), Epstein Barr virus, human
 CC immunodeficiency type-1 (HIV-1), human immunodeficiency type-2 (HIV-2),
 CC Papilloma virus, Lassa virus, Mycobacterium tuberculosis (MT), p53,
 CC murine p53 (mp53), CEA, HER2/neu, and tyrosine kinase related protein
 CC (TKP). The composition is useful for preventing or treating viral
 CC infections or cancer, such as prostate cancer, hepatitis B, hepatitis C,
 CC AIDS, renal carcinoma, cervical carcinoma, lymphoma, CMV or chondyroma

CC acuminatum. The composition is also be used for diagnosing such diseases.
 CC This sequence represents a peptide of the invention.

XX Sequence 9 AA;

Query Match 55.4%; Score 31; DB 8; Length 9;
 Best Local Similarity 60.0%; Pred. No. 1.7e+06;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 WLTW 8
 | : | :
 Db 5 WVTW 9

RESULT 12

AAB99759

ID AAB99759 standard; peptide; 10 AA.

XX AC AAB99759;

XX DT 21-SEP-2001 (first entry)

XX DE Rhesus D antibody binding peptide SEQ ID NO:4.

XX Rhesus D antibody binding peptide; Rhesus D; RhD; identification;
 KW anti-Rhesus D antibody; immunogen; epitope; diagnosis; therapy;
 KW prophylaxis; haemolytic disease of the newborn; HDN; ITP;
 KW idiopathic thrombocytopenic purpura; immunoglobulin.

XX Homo sapiens.

XX PN EP1106625-A1.

XX PD 13-JUN-2001.

XX PF 17-NOV-1999; 99EP-00122858.

XX PR 17-NOV-1999; 99EP-00122858.

XX PA (ZLBB-) ZLB BIOPLASMA AG.

XX PI Miescher S, Hofmann A, Fisch I;

XX WPI; 2001-383568/41.

XX Novel peptides capable of binding Rhesus D antibodies are used to
 CC manufacture an agent for the diagnosis, therapy or prophylaxis of
 CC diseases associated with Rhesus D antigen, e.g. hemolytic disease of the
 CC newborn (HDN).

XX Claim 1; Page 12; 19pp; English.

XX The present sequence represents a peptide capable of binding Rhesus D
 CC antibodies (I). Also described in the present invention are: (1) a
 CC nucleic acid (II) encoding (I); (2) a vector (III) comprising one or more
 CC (II) operably linked to an expression control system; (3) a cell (IV)
 CC comprising (II) or (III); (4) preparing (I); (5) identifying (M1)
 CC peptides having immunologic properties of Rhesus D protein epitopes
 CC comprising subjecting an antibody/antibody fragment recognising an
 CC epitope of Rhesus D protein to several panning rounds with a phage
 CC display library, and identifying immunogenic peptide sequences which are
 CC mimotopes which differ in their amino acid sequence from the amino acid
 CC sequences of Rhesus D protein; and (6) peptides (V) with immunological
 CC properties of Rhesus D protein epitopes obtained by (M1). (I) is used to
 CC manufacture an agent for the diagnosis, therapy or prophylaxis of
 CC diseases associated with Rhesus D antigen, e.g. haemolytic disease of the
 CC newborn (HDN) or idiopathic thrombocytopenic purpura (ITP), for the
 CC manufacture of an affinity reagent for anti-Rhesus D antibodies purified
 CC or removed from body fluids or immunoglobulin preparations. Using (I) as
 CC an immunogen to raise anti-Rhesus D antibodies avoids using immunisation
 CC with foreign erythrocytes thereby avoiding the risk of transmission of
 CC viral diseases like AIDS and hepatitis B

SQ Sequence 10 AA;
 Query Match 55.4%; Score 31; DB 4; Length 10;
 Best Local Similarity 44.4%; Pred. No. 3.5e+02;
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 GYWLTIWGX 10
 ||| |:
 DB 1 GYWSAKWAV 9

RESULT 13
 AAB99758
 ID AAB99758 standard; peptide; 10 AA.
 XX
 AC AAB99758;
 XX
 DT 21-SEP-2001 (first entry)
 XX
 DE Rhesus D antibody binding peptide SEQ ID NO:3.
 XX
 KW Rhesus D antibody binding peptide; Rhesus D; RHD; identification;
 KW anti-Rhesus D antibody; immunogen; epitope; diagnosis; therapy;
 KW prophylaxis; haemolytic disease of the newborn; HDN; ITP;
 KW idiopathic thrombocytopenic purpura; immunoglobulin.
 XX
 OS Homo sapiens.
 XX
 PN EP1106625-A1.
 XX
 PD 13-JUN-2001.
 XX
 PF 17-NOV-1999; 99EP-00122858.
 XX
 PR 17-NOV-1999; 99EP-00122858.
 XX
 PA (ZLBB-) ZLB BIOPLASMA AG.
 PI Miescher S, Hofmann A, Fisch I;
 XX
 DR WPI; 2001-383568/41.
 XX
 PT Novel peptides capable of binding Rhesus D antibodies are used to
 PT manufacture an agent for the diagnosis, therapy or prophylaxis of
 PT diseases associated with Rhesus D antigen, e.g. hemolytic disease of the
 PT newborn (HDN).
 XX
 PS Claim 1; Page 12; 19pp; English.
 CC The present sequence represents a peptide capable of binding Rhesus D
 CC antibodies (I). Also described in the present invention are: (1) a
 CC nucleic acid (II) encoding (I); (2) a vector (III) comprising one or more
 CC (II) operably linked to an expression control system; (3) a cell (IV)
 CC comprising (II) or (III); (4) preparing (I); (5) identifying (M1)
 CC peptides having immunologic properties of Rhesus D protein epitopes
 CC comprising subjecting an antibody/antibody fragment recognising an
 CC epitope of Rhesus D protein to several panning rounds with a phage
 CC display library, and identifying immunogenic peptide sequences which are
 CC mimotopes which differ in their amino acid sequence from the amino acid
 CC sequences of Rhesus D protein; and (6) peptides (V) with immunological
 CC properties of Rhesus D protein epitopes obtained by (M1). (I) is used to
 CC manufacture an agent for the diagnosis, therapy or prophylaxis of
 CC diseases associated with Rhesus D antigen, e.g. haemolytic disease of the
 CC newborn (HDN) or idiopathic thrombocytopenic purpura (ITP), for the
 CC manufacture of an affinity reagent for anti-Rhesus D antibodies purified
 CC or removed from body fluids or immunoglobulin preparations. Using (I) as
 CC an immunogen to raise anti-Rhesus D antibodies avoids using immunisation
 CC with foreign erythrocytes thereby avoiding the risk of transmission of
 CC viral diseases like AIDS and hepatitis B

SQ Sequence 10 AA;
 Query Match 55.4%; Score 31; DB 4; Length 10;
 Best Local Similarity 44.4%; Pred. No. 3.5e+02;
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 GYWLTIWGX 10
 ||| |:
 DB 1 GYWSAKWAV 9

RESULT 14
 ABP22594
 ID ABP22594 standard; peptide; 8 AA.
 XX
 AC ABP22594;
 XX
 DT 11-SEP-2003 (revised)
 DT 15-JUL-2002 (first entry)
 XX
 DE HIV A11 motif env peptide #317.
 XX
 KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
 KW vaccine; HIV infection; immunisation; virucide.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN WO200124810-A1.
 XX
 PD 12-APR-2001.
 XX
 PF 05-OCT-2000; 2000WO-US027766.
 XX
 PR 05-OCT-1999; 99US-00412863.
 XX
 PA (EPIM-) EPIMUNE INC.
 XX
 PI Sette A, Sidney J, Southwood S, Livingston BD, Ciesnut R;
 PI Baker DM, Celis E, Kubo RT, Grey HM;
 XX
 DR WPI; 2001-354887/37.
 XX
 PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
 PT peptide groups, useful for vaccinating against HIV-1.
 XX
 PS Claim 32; Page 332; 448pp; English.
 CC The present invention describes a composition (I) comprising a prepared
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
 CC be used for immunising subjects against HIV-1 infections. The use of
 CC group-based vaccines has several advantages over traditional vaccines,
 CC particularly when compared to the use of whole antigens in vaccine
 CC compositions. There is evidence that the immune response to whole
 CC antigens is directed largely toward variable regions of the antigen,
 CC allowing for immune escape due to mutations. The groups for inclusion in
 CC an group-based vaccine may be selected from conserved regions of viral or
 CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present
 CC in whole antigens can be avoided with the use of group-based vaccines. An
 CC additional advantage of an group-based vaccine approach is the ability to
 CC combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced
 CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response
 CC is not possible with traditional approaches. ABP1501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)

SQ Sequence 8 AA;
 Query Match 53.6%; Score 30; DB 4; Length 8;
 Best Local Similarity 66.7%; Pred. No. 1.7e+06;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LTIWGX 10
Db ||:|:|:
2 LTVWGI 7

Db ||:|:|:
2 LTVWGI 7

Search completed: January 3, 2005, 16:56:47
Job time : 149 secs

RESULT 15
ABP14287
ID ABP14287 standard; peptide; 8 AA.

XX AC ABP14287;

XX 11-SEP-2003 (revised)

DT 15-JUL-2002 (first entry)

XX HIV A03 super motif env peptide #27.

XX HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
KW vif; tat; cytototoxic T lymphocyte; CTL; immune response; epitope; antigen;
KW vaccine; HIV infection; immunisation; virucide.

XX Human immunodeficiency virus 1.

OS WO200124810-A1.

XX 12-APR-2001.

XX 05-OCT-2000; 2000WO-US027766.

XX 05-OCT-1999; 99US-00412863.

XX (EPIM-) EPIMMUNE INC.

XX Sette A, Sidney J, Southwood S, Livingston BD, Ciesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;

XX WPI; 2001-354887/37.

XX Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1.

XX Claim 32; Page 161; 448pp; English.

XX The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABL25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
CC be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,
CC allowing for immune escape due to mutations. The groups for inclusion in
CC an group-based vaccine may be selected from conserved regions of viral or
CC tumour-associated antigens, which therefore reduces the likelihood of
CC escape mutants. Furthermore, immunosuppressive groups that may be present
CC in whole antigens can be avoided with the use of group-based vaccines. An
CC additional advantage of an group-based vaccine approach is the ability to
CC combine selected groups (CTL and HTL), and further, to modify the
CC composition of the groups, achieving, for example, enhanced
CC immunogenicity. Accordingly, the immune response can be modulated, as
CC appropriate, for the target disease. Similar engineering of the response
CC is not possible with traditional approaches. ABP1501 to ABP25412
CC represent peptide sequences used in the exemplification of the present
CC invention. (Updated on 11-SEP-2003 to standardise OS field)

XX Sequence 8 AA;

Query Match 53.6%; Score 30; DB 4; Length 8;
Best Local Similarity 66.7%; Pred. NO. 1.7e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LTIWGX 10

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:50:00 ; Search time 37 Seconds
(without alignments)
17.924 Million cell updates/sec

Title: US-10-046-922-34

Perfect score: 56

Sequence: 1 XGYMLTWGX 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 110780

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pap: *
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pap: *
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pap: *
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pap: *
5: /cgn2_6/ptodata/1/1aa/PTUS-COMB.pap: *
6: /cgn2_6/ptodata/1/1aa/backfiles1.pap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	53.6	9	4	US-09-311-784A-348
2	30	53.6	10	2	US-08-735-253-8
3	30	53.6	10	2	US-08-735-253-13
4	29	51.8	7	4	US-09-069-827A-94
5	29	51.8	8	3	US-09-315-304B-1649
6	29	51.8	10	4	US-09-125-641-3
7	28	50.0	8	3	US-08-586-670A-17
8	28	50.0	8	3	US-09-082-279B-1495
9	28	50.0	8	4	US-09-834-784-1495
10	28	50.0	8	4	US-09-350-641C-1650
11	28	50.0	10	3	US-09-315-304B-1587
12	28	50.0	10	4	US-09-350-325-47
13	28	50.0	10	4	US-09-350-641C-1587
14	28	50.0	10	4	US-09-620-091-28
15	27.5	49.1	6	1	US-08-191-571-12
16	27.5	49.1	6	5	PCT-US95-00296-12
17	27	48.2	8	3	US-09-082-279B-1500
18	27	48.2	8	4	US-09-834-784-1500
19	27	48.2	8	4	US-09-350-641C-1655
20	27	48.2	9	1	US-08-024-253-18
21	27	48.2	10	4	US-09-239-043D-2474
22	26	46.4	5	1	US-08-353-400-27
23	26	46.4	6	1	US-08-415-099A-3
24	26	46.4	8	1	US-08-415-099A-4
25	26	46.4	8	3	US-08-467-472C-5
26	26	46.4	8	3	US-08-467-472C-6
27	26	46.4	8	3	US-09-384-061-5

28	26	46.4	8	3	US-09-384-061-6	Sequence 6, Appli
29	26	46.4	8	4	US-09-125-641-1	Sequence 1, Appli
30	26	46.4	8	4	US-09-125-641-2	Sequence 2, Appli
31	26	46.4	8	4	US-08-632-444B-9	Sequence 9, Appli
32	26	46.4	8	4	US-09-852-870A-5	Sequence 5, Appli
33	26	46.4	9	3	US-08-985-526-13	Sequence 13, Appli
34	26	46.4	9	4	US-09-125-641-29	Sequence 29, Appli
35	26	46.4	10	1	US-08-208-886C-83	Sequence 83, Appli
36	26	46.4	10	1	US-08-208-886C-87	Sequence 87, Appli
37	26	46.4	10	1	US-08-704-744-85	Sequence 85, Appli
38	26	46.4	10	1	US-08-704-744-89	Sequence 89, Appli
39	26	46.4	10	1	US-08-469-557-64	Sequence 64, Appli
40	26	46.4	10	1	US-08-469-557-68	Sequence 68, Appli
41	26	46.4	10	2	US-08-290-793B-64	Sequence 64, Appli
42	26	46.4	10	2	US-08-290-793B-68	Sequence 68, Appli
43	26	46.4	10	3	US-09-227-357-377	Sequence 377, App
44	26	46.4	10	4	US-09-125-641-19	Sequence 19, Appli
45	26	46.4	10	4	US-08-632-444B-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1

US-09-311-784A-348
; Sequence 348, Application US/09311784A

; Patent No. 6534482

; GENERAL INFORMATION:

; APPLICANT: Fikes, John D.

; APPLICANT: Hermanson, Gary G.

; APPLICANT: Sette, Alessandro

; APPLICANT: Ishioka, Glenn Y.

; APPLICANT: Livingston, Brian

; APPLICANT: Chesnut, Robert W.

; APPLICANT: Epimmune Inc.

; TITLE OF INVENTION: Expression Vectors for Stimulating an Immune Response and Methods of Using the Same

; FILE REFERENCE: 39963-20022.01

; CURRENT APPLICATION NUMBER: US/09/311,784A

; CURRENT FILING DATE: 1999-05-13

; PRIOR APPLICATION NUMBER: US 60/085,751

; PRIOR FILING DATE: 1998-05-15

; NUMBER OF SEQ ID NOS: 463

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 348

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: HIV1 ENV 61 (peptide 25.0032)

US-09-311-784A-348

Query Match 53.6%; Score 30; DB 4; Length 9;

Best Local Similarity 66.7%; Pred. No. 3.8e+05;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LTIMGX 10

DB 4 LTVWGI 9

RESULT 2

US-08-735-253-8

; Sequence 8, Application US/08735253

; Patent No. 5942491

; GENERAL INFORMATION:

; APPLICANT: Root-Bernstein, Robert S.

; TITLE OF INVENTION: Methods and Compositions for Treating

; TITLE OF INVENTION: Arthritis

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESS: Dressler, Goldsmith, Milnamow & Katz, Ltd.

; STREET: 180 N. Stetson

;
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/735,253
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5942491thrup, Thomas E.
; REGISTRATION NUMBER: 33,268
; REFERENCE/DOCKET NUMBER: MIC3302P0010US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5400
; TELEFAX: (312) 616-5460
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-735-253-8

Query Match 53.6%; Score 30; DB 2; Length 10;
Best Local Similarity 37.5%; Pred. No. 1.5e+02;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 YWLTWGX 10
; : : : :
; 1 FWRFLWGS 8
Db

RESULT 3
US-08-735-253-13
; Sequence 13, Application US/08735253
; Patent No. 5942491
; GENERAL INFORMATION:
; APPLICANT: Root-Bernstein, Robert S.
; TITLE OF INVENTION: Methods and Compositions for Treating
; TITLE OF INVENTION: Arthritis
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Milnamow & Katz, Ltd.
; STREET: 180 N. Stetson
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/735,253
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5942491thrup, Thomas E.
; REGISTRATION NUMBER: 33,268
; REFERENCE/DOCKET NUMBER: MIC3302P0010US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5400
; TELEFAX: (312) 616-5460
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids

;
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-735-253-13

Query Match 53.6%; Score 30; DB 2; Length 10;
Best Local Similarity 37.5%; Pred. No. 1.5e+02;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 YWLTWGX 10
; : : : :
; 1 FWRFLWGS 8
Db

RESULT 4
US-09-069-827A-94
; Sequence 94, Application US/09069827A
; Patent No. 6617114
; GENERAL INFORMATION:
; APPLICANT: FOWLKES, Dana M
; KAY, Brian K
; FRELINGER, Jeffrey A
; HYDE-DERUYSCHE, Robin P
; TITLE OF INVENTION: IDENTIFICATION OF DRUGS USING
; COMPLEMENTARY COMBINATORIAL LIBRARIES
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/069,827A
; FILING DATE: 30-Apr-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/050,359
; FILING DATE: 31-MAR-1998
; APPLICATION NUMBER: PCT/US97/19638
; FILING DATE: 31-OCT-1997
; APPLICATION NUMBER: US 08/740,671
; FILING DATE: 31-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Iver P
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: FOWLKES=4C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 94:
US-09-069-827A-94

Query Match 51.8%; Score 29; DB 4; Length 7;
Best Local Similarity 57.1%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 YWLTWGX 9
; : : : :
; 1 FWRFLWGS 8
Db

Db 1 YWPDWG 7

RESULT 5

US-09-315-304B-1649
; Sequence 1649, Application US/09315304B
; Patent No. 6348568
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merucka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-052
; CURRENT APPLICATION NUMBER: US/09/315,304B
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1667
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1649
; LENGTH: 8
; TYPE: PRT
; ORGANISM: HIV-1
US-09-315-304B-1649

Query Match 51.8%; Score 29; DB 3; Length 8;
Best Local Similarity 42.9%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 WLTWGX 10
| : : | :
Db 1 WNSLWG 7

RESULT 6

US-09-125-641-3
; Sequence 3, Application US/09125641
; Patent No. 6610297
; GENERAL INFORMATION:
; APPLICANT: Kricek, Franz
; APPLICANT: Stadler, Boda
; TITLE OF INVENTION: Peptide Immunogens For Vaccination
; TITLE OF INVENTION: Against and treatment of Allergy
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6610297artis Corporation
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/125,641
; FILING DATE: 21-AUG-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP97/01013
; FILING DATE: 28-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9604412.8
; FILING DATE: 01-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9617702.7
; FILING DATE: 22-AUG-1996
; ATTORNEY/AGENT INFORMATION:

NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 4-900-9862/A/NFI/PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 522-6923
TELEFAX: (908) 522-6923
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-09-125-641-3

Query Match 51.8%; Score 29; DB 4; Length 10;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYWL 5
: | | | |
Db 5 GGYWL 9

RESULT 7

US-08-586-670A-17
; Sequence 17, Application US/08586670A
; Patent No. 6241965
; GENERAL INFORMATION:
; APPLICANT: McBride, William
; APPLICANT: Dean, Richard T.
; TITLE OF INVENTION: Somatostatin Derivatives
; TITLE OF INVENTION: And their Radiolabeled Products
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/586,670A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6241965nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 92,385-DD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1..2
OTHER INFORMATION: /label= Variant residues
OTHER INFORMATION: /note= "Phe is in the D conformation and is
OTHER INFORMATION: linked to DTPA;

```
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..4
; OTHER INFORMATION: /label= Variant residues
; OTHER INFORMATION: /note= "The phe is in the D conformation; Xaa
; OTHER INFORMATION: is L-4-chlorophenylalanine; the Trp is in the
; OTHER INFORMATION: D conformation;
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 7..8
; OTHER INFORMATION: /label= Variant residues
; OTHER INFORMATION: /note= "The carboxyl group of the C-terminal
; OTHER INFORMATION: Thr is reduced to an alcohol;
US-08-586-670A-17

Query Match          50.0%; Score 28; DB 3; Length 8;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWLT 6
Db 1 FGWYKT 6

RESULT 8
US-09-082-279B-1495
; Sequence 1495, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohmed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1495
; LENGTH: 8
; TYPE: PRT
; ORGANISM: HIV-1
US-09-082-279B-1495

Query Match          50.0%; Score 28; DB 3; Length 8;
Best Local Similarity 42.9%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 WLTIWGX 10
Db 1 WDSLWGW 7

RESULT 9
US-09-834-784-1495
; Sequence 1495, Application US/09834784
; Patent No. 6562787
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohmed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/834,784
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
```

```
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1495
; LENGTH: 8
; TYPE: PRT
; ORGANISM: HIV-1
US-09-834-784-1495

Query Match          50.0%; Score 28; DB 4; Length 8;
Best Local Similarity 42.9%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 WLTIWGX 10
Db 1 WDSLWGW 7

RESULT 10
US-09-350-641C-1650
; Sequence 1650, Application US/09350641C
; Patent No. 6856906
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-067
; CURRENT APPLICATION NUMBER: US/09/350,641C
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1650
; LENGTH: 8
; TYPE: PRT
; ORGANISM: HIV-1
US-09-350-641C-1650

Query Match          50.0%; Score 28; DB 4; Length 8;
Best Local Similarity 42.9%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 WLTIWGX 10
Db 1 WDSLWGW 7

RESULT 11
US-09-315-304B-1587
; Sequence 1587, Application US/09315304B
; Patent No. 6348588
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-052
; CURRENT APPLICATION NUMBER: US/09/315,304B
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1667
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1587
```

```
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-315-304B-1587

Query Match          50.0%; Score 28; DB 3; Length 10;
Best Local Similarity 33.3%; Pred. No. 2.9e+02;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      2 GYWLTIWG 10
Db      1 GGWASLWNW 9

RESULT 12
US-09-350-325-47
; Sequence 47, Application US/09350325
; Patent No. 6541020
; GENERAL INFORMATION:
; APPLICANT: Ding, S.
; APPLICANT: Kang, M.
; APPLICANT: Venetta, T.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ADMINISTRATION OF
; TITLE OF INVENTION: THERAPEUTIC REAGENTS
; FILE REFERENCE: 7872-062
; CURRENT APPLICATION NUMBER: US/09/350,325
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 47
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: enhancer peptide
US-09-350-325-47

Query Match          50.0%; Score 28; DB 4; Length 10;
Best Local Similarity 33.3%; Pred. No. 2.9e+02;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      2 GYWLTIWG 10
Db      1 GGWASLWNW 9

RESULT 13
US-09-350-641C-1587
; Sequence 1587, Application US/09350641C
; Patent No. 6656906
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M. D.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; TITLE OF INVENTION: PROPERTIES
; FILE REFERENCE: 7872-067
; CURRENT APPLICATION NUMBER: US/09/350,641C
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1587
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-350-641C-1587

Query Match          50.0%; Score 28; DB 4; Length 10;
Best Local Similarity 33.3%; Pred. No. 2.9e+02;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      2 GYWLTIWG 10
Db      1 GGWASLWNW 9

RESULT 14
US-09-620-091-28
; Sequence 28, Application US/09620091
; Patent No. 6716811
; GENERAL INFORMATION:
; APPLICANT: CWIRLA, STEVEN E.
; APPLICANT: BALU, PALANI
; APPLICANT: DUFFIN, DAVID J.
; APPLICANT: PIPLANI, SUNILA
; APPLICANT: MERRILL, BARBARA MCBOWEN
; APPLICANT: SCHATZ, PETER JOSEPH
; TITLE OF INVENTION: COMPOUNDS HAVING AFFINITY FOR THE GRANULOCYTE-COLONY
; TITLE OF INVENTION: STIMULATING FACTOR RECEPTOR (G-CSFR) AND ASSOCIATED
; TITLE OF INVENTION: USES
; FILE REFERENCE: 0300-0014
; CURRENT APPLICATION NUMBER: US/09/620,091
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 491
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-620-091-28

Query Match          50.0%; Score 28; DB 4; Length 10;
Best Local Similarity 25.0%; Pred. No. 2.9e+02;
Matches 2; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      1 XGYWLTIW 8
Db      2 ESFWVELW 9

RESULT 15
US-08-191-571-12
; Sequence 12, Application US/08191571
; Patent No. 5521156
; GENERAL INFORMATION:
; APPLICANT: Owen, Thomas J.
; APPLICANT: Kudlacz, Elizabeth M
; APPLICANT: Buck, Stephen H
; APPLICANT: Harbeson, Scott L
; TITLE OF INVENTION: Cyclic Neurokinin A Antagonists
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Marion Merrell Dow Inc.
; STREET: P. O. Box 156300 2110 E. Galbraith Rd.
; CITY: Cincinnati
; STATE: Ohio
; COUNTRY: United States
; ZIP: 45215-6300
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/191,571
; FILING DATE: 03-FEB-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Boudreaux, William R
; REGISTRATION NUMBER: 35,796
; REFERENCE/DOCKET NUMBER: M01718
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 513-948-6566
; TELEFAX: 513-948-7961 or 4681
; TELEX: 214320
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; MOLECULE TYPE: peptide
US-08-191-571-12

Query Match      49.1%; Score 27.5; DB 1; Length 6;
Best Local Similarity 71.4%; Pred. No. 3.8e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY      3 YWLTWG 9
Db      1 YWLR-WG 6

Search completed: January 3, 2005, 17:01:24
Job time : 37 secs

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:57:37 ; Search time 139 Seconds
(without alignments)
25.880 Million cell updates/sec

Title: US-10-046-922-34
Perfect score: 56
Sequence: 1 XGYWLTWGX 10

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues
Total number of hits satisfying chosen parameters: 182644

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	100.0	10	13	US-10-046-922-34
2	56	100.0	10	13	US-10-046-922-35
3	32	57.1	9	13	US-10-046-922-36
4	32	57.1	10	10	US-09-572-404B-1454
5	30	53.6	9	14	US-09-894-018-154
6	30	53.6	9	14	US-10-133-210-8
7	30	53.6	9	14	US-10-371-069-348
8	30	53.6	9	14	US-10-371-069-348
9	30	53.6	9	14	US-10-371-069-348
10	30	53.6	9	14	US-10-371-069-348
11	30	53.6	9	15	US-10-182-252A-180
12	30	53.6	9	15	US-10-182-252A-181
13	30	53.6	9	15	US-10-182-252A-305

14	30	53.6	9	15	US-10-182-252A-792	Sequence 792, App
15	30	53.6	9	15	US-10-182-252A-793	Sequence 793, App
16	30	53.6	9	15	US-10-182-252A-837	Sequence 837, App
17	30	53.6	9	15	US-10-182-252A-838	Sequence 838, App
18	30	53.6	9	15	US-10-182-252A-1271	Sequence 1271, App
19	30	53.6	9	17	US-10-474-960A-154	Sequence 154, App
20	30	53.6	10	8	US-08-821-739A-92	Sequence 92, Appl
21	30	53.6	10	10	US-09-793-451-69	Sequence 69, Appl
22	30	53.6	10	14	US-10-133-210-7	Sequence 7, Appl
23	30	53.6	10	14	US-10-133-210-40	Sequence 40, Appl
24	30	53.6	10	14	US-10-133-210-70	Sequence 70, Appl
25	30	53.6	10	14	US-10-200-708-581	Sequence 581, App
26	30	53.6	10	14	US-10-283-722-69	Sequence 69, Appl
27	30	53.6	10	14	US-10-283-903-69	Sequence 707, Appl
28	30	53.6	10	15	US-10-462-452-707	Sequence 812, App
29	30	53.6	10	15	US-10-601-953-812	Sequence 708, App
30	30	53.6	10	16	US-10-322-862-708	Sequence 181, App
31	29	51.8	10	14	US-10-050-902-181	Sequence 181, App
32	29	51.8	10	14	US-10-050-898-181	Sequence 181, App
33	29	51.8	10	14	US-10-200-708-355	Sequence 355, App
34	29	51.8	10	15	US-10-617-876-33	Sequence 33, Appl
35	28	50.0	8	14	US-10-351-641-1650	Sequence 1650, Ap
36	28	50.0	10	14	US-10-347-562-47	Sequence 47, Appl
37	28	50.0	10	14	US-10-351-641-1587	Sequence 1587, Ap
38	28	50.0	10	15	US-10-239-656-34	Sequence 34, Appl
39	28	50.0	10	15	US-10-239-656-44	Sequence 44, Appl
40	28	50.0	10	16	US-10-327-598-450	Sequence 450, App
41	27	48.2	8	14	US-10-351-641-1655	Sequence 1655, Ap
42	27	48.2	9	10	US-09-809-638-145	Sequence 145, App
43	27	48.2	9	10	US-09-809-638-246	Sequence 246, App
44	27	48.2	9	10	US-09-809-638-417	Sequence 417, App
45	27	48.2	9	10	US-09-809-638-523	Sequence 523, App

ALIGNMENTS

RESULT 1

US-10-046-922-34
; Sequence 34, Application US/10046922
; Publication NO. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Brkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 10
; TYPE: PRT
; ORGANISM: isolated peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(1)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (10)..(10)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-34

Query Match 100.0%; Score 56; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYWLTWGX 10

DB 1 XGYWLTWGX 10

```
RESULT 2
US-10-046-922-35
; Sequence 35, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35
; LENGTH: 10
; TYPE: PRT
; ORGANISM: isolated peptide
US-10-046-922-35

Query Match      100.0%; Score 56; DB 13; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.18;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 XGYWLTWGX 10
Db      1 CGYWLTIWGC 10
      :|||||:
      :|||||:

RESULT 3
US-10-046-922-36
; Sequence 36, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 9
; TYPE: PRT
; ORGANISM: peptide
US-10-046-922-36

Query Match      57.1%; Score 32; DB 13; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.5e+06;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 XGYWLTW 8
Db      1 SGYWDWTW 8
      :|||
      :|||

RESULT 4
US-09-572-404B-1454
; Sequence 1454, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 1454
; LENGTH: 10
; TYPE: PRT
```

```
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in P2RY5 at 139-148 and may interact with Sequence
; OTHER INFORMATION: in this patent.
US-09-572-404B-1454

Query Match      57.1%; Score 32; DB 10; Length 10;
Best Local Similarity 55.6%; Pred. No. 4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GYWLTIWGX 10
Db      1 GWLTVIGG 9
      :|||:
      :|||:

RESULT 5
US-09-894-018-154
; Sequence 154, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 154
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Transgenic mouse
US-09-894-018-154

Query Match      53.6%; Score 30; DB 9; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      5 LTIWGX 10
Db      4 LTVWGI 9
      :|||:
      :|||:

RESULT 6
US-10-133-210-8
; Sequence 8, Application US/10133210
; Publication No. US20030103964A1
; GENERAL INFORMATION:
; APPLICANT: DeLisi, Charles
; APPLICANT: Berzofsky, Jay
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Vaccaro, Dennis
; APPLICANT: Weng, Zhiping
; APPLICANT: Zhang, Chao
; TITLE OF INVENTION: METHODS FOR DESIGNING MOLECULAR CONJUGATES AND
; FILE REFERENCE: BU-03SAX
; CURRENT APPLICATION NUMBER: US/10/133,210
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 281
; SOFTWARE: PatentIn Ver. 2.0
```

; SEQ ID NO 8
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-133-210-8

Query Match 53.6%; Score 30; DB 14; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LTIWGX 10
||:|:
Db 4 LTVWGI 9

RESULT 7
US-10-371-525-348
; Sequence 348, Application US/10371525
; Publication No. US20030203869A1
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/10/371,525
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/311,784
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HIV1 ENV 61 (peptide 25.0032)
US-10-371-525-348

Query Match 53.6%; Score 30; DB 14; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LTIWGX 10
||:|:
Db 4 LTVWGI 9

RESULT 8
US-10-371-069-348
; Sequence 348, Application US/10371069
; Publication No. US20030216342A1
; GENERAL INFORMATION:
; APPLICANT: Epimmune Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.10

; CURRENT APPLICATION NUMBER: US/10/371,069
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HIV1 ENV 61 (peptide 25.0032)
US-10-371-069-348

Query Match 53.6%; Score 30; DB 14; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LTIWGX 10
||:|:
Db 4 LTVWGI 9

RESULT 9
US-10-371-645-348
; Sequence 348, Application US/10371645
; Publication No. US20030216343A1
; GENERAL INFORMATION:
; APPLICANT: Epimmune Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.11
; CURRENT APPLICATION NUMBER: US/10/371,645
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HIV1 ENV 61 (peptide 25.0032)
US-10-371-645-348

Query Match 53.6%; Score 30; DB 14; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LTIWGX 10
||:|:
Db 4 LTVWGI 9

RESULT 10
US-10-371-260-348
; Sequence 348, Application US/10371260
; Publication No. US20030220285A1
; GENERAL INFORMATION:
; APPLICANT: Epimmune Inc.
; APPLICANT: Fikes, John D.

US-10-182-252A-305

Query Match 53.6%; Score 30; DB 15; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LTIWGX 10
|||:
Db 4 LTVWGV 9

RESULT 14

US-10-182-252A-792
; Sequence 792, Application US/10182252A
; Publication No. US20040072162A1
; GENERAL INFORMATION:
; APPLICANT: FOMSGAARD, ANDERS
; APPLICANT: BRUNAK, SOREN
; APPLICANT: BUUS, SOREN
; APPLICANT: CORBET, SYLVIE
; APPLICANT: LAUEMOLLER, SANNE LISE
; APPLICANT: HANSEN, JAN
; TITLE OF INVENTION: HIV PEPTIDE AND NUCLEIC ACIDS ENCODING THEM FOR DIAGNOSIS AND
; TITLE OF INVENTION: CONTROL OF HIV INFECTIONS
; FILE REFERENCE: 030307/0205
; CURRENT APPLICATION NUMBER: US/10/182,252A
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: PCT/DK01/00059
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: EP 00610017.6
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/179,333
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 1388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 792
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV peptide
US-10-182-252A-792

Query Match 53.6%; Score 30; DB 15; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LTIWGX 10
|||:
Db 4 LTVWGA 9

RESULT 15

US-10-182-252A-793
; Sequence 793, Application US/10182252A
; Publication No. US20040072162A1
; GENERAL INFORMATION:
; APPLICANT: FOMSGAARD, ANDERS
; APPLICANT: BRUNAK, SOREN
; APPLICANT: BUUS, SOREN
; APPLICANT: CORBET, SYLVIE
; APPLICANT: LAUEMOLLER, SANNE LISE
; APPLICANT: HANSEN, JAN
; TITLE OF INVENTION: HIV PEPTIDE AND NUCLEIC ACIDS ENCODING THEM FOR DIAGNOSIS AND
; TITLE OF INVENTION: CONTROL OF HIV INFECTIONS
; FILE REFERENCE: 030307/0205
; CURRENT APPLICATION NUMBER: US/10/182,252A
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: PCT/DK01/00059
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: EP 00610017.6
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/179,333

; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 1388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 793
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV peptide
US-10-182-252A-793

Query Match 53.6%; Score 30; DB 15; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LTIWGX 10
|||:
Db 4 LTVWGL 9

Search completed: January 3, 2005, 17:11:57
Job time : 140 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:35:00 ; Search time 38 seconds
(without alignments)
25.320 Million cell updates/sec

Title: US-10-046-922-34

Perfect score: 56

Sequence: 1 XGYWLTWGX 10

Scoring table: <H20SUM62DX>
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	44.6	10	2 S71868	glutathione transf
2	23	41.1	7	2 PT0586	T-cell receptor be
3	22	39.3	10	2 T17075	cytochrome-c oxida
4	20	35.7	5	2 JH0253	gut pentapeptide -
5	20	35.7	9	2 PT0324	Ig heavy chain CRD
6	20	35.7	10	2 PT0289	Ig heavy chain CRD
7	19	33.9	10	2 PT0230	Ig heavy chain CDR
8	18	32.1	4	2 B53284	T-cell receptor be
9	18	32.1	7	2 PT0642	T-cell receptor be
10	18	32.1	7	2 PT0728	T-cell receptor be
11	18	32.1	7	2 B48394	major fat-globule
12	18	32.1	7	2 S57274	triacylglycerol li
13	18	32.1	8	2 JS0315	leucokinin V - Mad
14	18	32.1	8	2 A31570	angiotensin-conver
15	18	32.1	9	2 PT0634	T-cell receptor be
16	18	32.1	9	2 PT0562	T-cell receptor be
17	18	32.1	9	2 A60522	sperm-activating p
18	18	32.1	10	2 B33995	hypotrehalosemic h
19	18	32.1	10	2 S08997	hypotrehalosemic
20	18	32.1	10	2 A60421	hypotrehalosemic
21	18	32.1	10	2 S08998	hypotrehalosemic
22	18	32.1	10	2 A26381	hypotrehalosemic
23	18	32.1	10	2 JC1416	hypotrehalosemic
24	18	32.1	10	2 S09138	hypotrehalosemic
25	18	32.1	10	2 A31571	hypotrehalosemic/
26	18	32.1	10	2 PH1344	Ig heavy chain DJ
27	18	32.1	10	2 PH0923	T-cell receptor be
28	18	32.1	10	2 A40753	aldehyde ferredoxi
29	18	32.1	10	2 H37196	bradykinin-potent

30	18	32.1	10	2 F33932	Ig mu chain J regi
31	18	32.1	10	2 S53789	neuropeptide Pec-H
32	18	32.1	10	2 A59173	nuclease Bhl (EC 3
33	17	30.4	6	2 PT0629	T-cell receptor be
34	17	30.4	6	2 PT0637	T-cell receptor be
35	17	30.4	6	2 A61068	locustakinin - mig
36	17	30.4	7	2 PT0628	T-cell receptor be
37	17	30.4	7	2 PT0722	T-cell receptor be
38	17	30.4	7	2 PX0008	glucuronosyltransf
39	17	30.4	7	2 PD0029	pev-kinin 1 - pena
40	17	30.4	7	2 S33244	neuromodulatory pe
41	17	30.4	7	2 S33245	neuromodulatory pe
42	17	30.4	7	2 S33246	neuromodulatory pe
43	17	30.4	8	2 PT0724	T-cell receptor be
44	17	30.4	8	2 JS0316	leucokinin VI - Ma
45	17	30.4	8	2 JS0317	leucokinin VII - M

ALIGNMENTS

RESULT 1

S71868

glutathione transferase (EC 2.5.1.18) class mu 4 - pig (fragment)

N;Alternate names: glutathione S-transferase class mu 4

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 19-Mar-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004

C;Accession: S71868

R;Routini, P.; Anglade, P.; Debrauwer, L.; Tulliez, J.

Biochem. J. 317, 879-884, 1996

A;Title: Characterization of pig liver glutathione S-transferases using HPLC-electrospra

A;Reference number: S71864; MUID:96332484; PMID:8760377

A;Accession: S71868

A;Molecule type: protein

A;Residues: 1-10 <ROU>

A;Cross-references: UNIPROT:Q7M3E8

C;Comment: At least five species-independent classes of cytosolic glutathione transferase

s mitochondrial form are known.

C;Complex: dimer

C;Function:

A;Description: catalyzes the nucleophilic conjugation of intracellular glutathione to a

A;Pathway: detoxification; xenobiotics metabolism

A;Note: increased hydrophilicity of GSH-conjugates facilitates their further metabolism

es of damage

C;Superfamily: glutathione transferase

C;Keywords: dimer; transferase

Query Match 44.6%; Score 25; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 4.8e+02;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYW 4

DB 3 LGYW 6

RESULT 2

PT0586

T-cell receptor beta chain V-D-J region (141-1CN) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C;Accession: PT0586; PT0592

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0586

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-7 <FE>

C;Experimental source: day 19 fetal thymus, strain BALB/c (clones 141-1CN and 141-1CD)

C;Keywords: T-cell receptor

Query Match 41.1%; Score 23; DB 2; Length 7;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 LTWGX 10
:|:|:
Db 3 SIWGG 7

RESULT 3
T17075
cytochrome-c oxidase (EC 1.9.3.1) chain I - Chamaeleo fischeri mitochondrion (fragment)
C;Species: mitochondrion Chamaeleo fischeri
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17075
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Papenfuss, T.J.
J. Mol. Evol. 44, 660-674, 1997
A;Title: Evolutionary shifts in three major structural features of the mitochondrial gene
A;Reference number: Z18674; MUID:97315309; PMID:19169559
A;Accession: T17075
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-10 <MAC>
A;Cross-references: UNIPROT:O79912; EMBL:U82688; NID:g3603112; PID:g3603115; PIDN:AAC622
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

Query Match 39.3%; Score 22; DB 2; Length 10;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 WLTIW 8
:|:|:
Db 2 WLLRW 6

RESULT 4
JH0253
gut pentapeptide - Japanese eel
C;Species: Anguilla japonica (Japanese eel)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Apr-1995
C;Accession: JH0253
R;Uesaka, T.; Ikeda, T.; Kubota, I.; Muneoka, Y.; Ando, M.
Biochem. Biophys. Res. Commun. 180, 828-832, 1991
A;Title: Structure and function of a pentapeptide isolated from the gut of the eel.
A;Reference number: JH0253; MUID:92062113; PMID:1953755
A;Accession: JH0253
A;Molecule type: protein
A;Residues: 1-5 <UES>
A;Experimental source: gut
C;Comment: This peptide increased basal tone of the circular muscle of the esophagogastric
, and of the circular muscle of the gastro-intestinal junction.

Query Match 35.7%; Score 20; DB 2; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GYW 4
:|:|:
Db 1 GFW 3

RESULT 5
PT0324
Ig heavy chain CDR3 region (clone J2-106C) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0324
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j

A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0324
A;Molecule type: DNA
A;Residues: 1-9 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 35.7%; Score 20; DB 2; Length 9;
Best Local Similarity 37.5%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 XGYWLTW 8
:|:|:|:
Db 2 PGYGESTW 9

RESULT 6
PT0289
Ig heavy chain CDR3 region (clone 4-109) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0289
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0289
A;Molecule type: DNA
A;Residues: 1-10 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 35.7%; Score 20; DB 2; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.7e+03;
Matches 2; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 WLTIWGX 10
:|:|:|:
Db 4 WISWGGG 10

RESULT 7
PT0230
Ig heavy chain CDR3 region (clone 1-118A) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0230
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0230
A;Molecule type: DNA
A;Residues: 1-10 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 33.9%; Score 19; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 3.8e+03;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 LTIWGX 10
:|:|:|:
Db 3 ITIFGV 8

RESULT 8
B53284
T-cell receptor beta 2 chain D region, Dbeta2 - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C;Accession: B53284
R;Harindranath, N.; Alexander, C.B.; Mage, R.G.

Mol. Immunol. 28, 881-888, 1991
A:Title: Evolutionarily conserved organization and sequences of germline diversity and
A:Reference number: A53284; MUID:91342695; PMID:1678859
A:Accession: B53284
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-4 <HAR>
A:CROSS-references: GB:S60737; NID:G233916; PIDN:AA19518.1; PID:G233918
A:Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBI:P:60738)
C:Keywords: T-cell receptor

Query Match 32.1%; Score 18; DB 2; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 WGX 10
||:
Db 2 WGG 4

RESULT 9

PT0642
T-cell receptor beta chain V-D-J region (111-1H) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
A:Accession: PT0642
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0642
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-7 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 32.1%; Score 18; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 WGX 10
||:
Db 5 WGG 7

RESULT 10

PT0728
T-cell receptor beta chain V-D-J region (161-2H) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
A:Accession: PT0728
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0728
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-7 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 32.1%; Score 18; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 WGX 10
||:
Db 5 WGG 7

RESULT 11

B48394

major fat-globule membrane protein Gp 55 - guinea pig (fragment)
C:Species: Cavia porcellus (guinea pig)
C:Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 23-Mar-1995
A:Accession: B48394
R:Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
II-like sequences.
A:Reference number: A48394; MUID:93250576; PMID:8485470
A:Accession: B48394
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <MAT>
A:Experimental source: milk
A:Note: sequence extracted from NCBI backbone (NCBI:P:131444)

Query Match 32.1%; Score 18; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 WGX 10
||:
Db 1 WGP 3

RESULT 12

S57274
triacylglycerol lipase (EC 3.1.1.3) - Psychrobacter immobilis (fragment)
C:Species: Psychrobacter immobilis
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
A:Accession: S57274
R:Arpigny, J.L.; Feller, G.; Gerday, C.
Biochim. Biophys. Acta 1263, 103, 1995
A:Title: Corrigendum to "Cloning, sequence and structural features of a lipase from the
A:Reference number: S57274; MUID:95359197; PMID:7632728
A:Accession: S57274
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-7 <ARP>
A:CROSS-references: UNIPROT:Q02104; EMBL:X67712
C:Keywords: carboxylic ester hydrolase

Query Match 32.1%; Score 18; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 WGX 10
||:
Db 1 WGD 3

RESULT 13

JS0315
leucokinin V - Madeira cockroach
C:Species: Leucophaea madeira (Madeira cockroach)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
A:Accession: JS0315
R:Holman, G.M.; Cook, B.J.; Nachman, R.J.
Comp. Biochem. Physiol. C 88, 27-30, 1987
A:Title: Isolation, primary structure, and synthesis of leucokinins V and VI: myotropic
A:Reference number: JS0315
A:Accession: JS0315
A:Molecule type: protein
A:Residues: 1-8 <HOL>
A:CROSS-references: UNIPROT:P19987
C:Comment: Leucokinins, a family of cephalomyotropic peptides, stimulate contractile acti
C:Keywords: amidated carboxyl end; cephalomyotropic peptide
P:8/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 32.1%; Score 18; DB 2; Length 8;
Best Local Similarity 37.5%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 GYWLITWG 9
Db 1 GSGFSSWG 8

RESULT 14

A31570
angiotensin-converting enzyme inhibitor - yellowfin tuna
C:Species: Thunnus albacares (yellowfin tuna)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C:Accession: A31570
R:Kohama, Y.; Matsumoto, S.; Oka, H.; Teramoto, T.; Okabe, M.; Mimura, T.
Biochem. Biophys. Res. Commun. 155, 332-337, 1988
A:Title: Isolation of angiotensin-converting enzyme inhibitor from tuna muscle.
A:Reference number: A31570; PMID:88326322; PMID:3415688
A:Accession: A31570
A:Molecule type: protein
A:Residues: 1-8 <KOH>
A:Cross-References: UNIPROT:P18691
A:Note: the source is designated as Neothunnus macropterus
C:Superfamily: unassigned animal peptides
C:Keywords: angiotensin-converting enzyme inhibitor

Query Match 32.1%; Score 18; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
Matches 2; Conservative 1; Mismatches 0;

QY 8 WGX 10
Db 6 WGD 8

RESULT 15

PT0634
T-cell receptor beta chain V-D-J region (121-2CL) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0634
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; PMID:91277601; PMID:1711558
A:Accession: PT0634
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-9 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 32.1%; Score 18; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 WGX 10
Db 6 WGG 8

Search completed: January 3, 2005, 16:57:29
Job time : 38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:49:20 ; Search time 185 Seconds
(without alignments)
31.101 Million cell updates/sec

Title: US-10-046-922-34
Perfect score: 56
Sequence: 1 XGYWLTWGX 10

Scoring table: 4BLOSUM62DX**4
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 2971

Minimum DB seq length: 0
Maximum DB seq length: 1000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02:
1: uniprot_sprot:
2: uniprot_trembl:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32.5	58.0	10	1 LABA_JATMU	P13270 jatropa mu
2	27	48.2	8	2 Q70Y57	Q70Y57 fuerstia af
3	27	48.2	8	2 CAD45547	Cad45547 fuerstia
4	27	48.2	10	2 Q8SHF6	Q8SHF6 chamaeleo m
5	25	44.6	10	2 Q7M3E8	Q7M3E8 sus scrofa
6	25	44.6	10	2 Q6JL97	Q6JL97 neisseria g
7	25	44.6	10	2 AAS16521	Aas16521 neisseria
8	24.5	43.8	9	2 Q85DB0	Q85DB0 lepitemur s
9	24.5	43.8	9	2 Q85DB8	Q85DB8 lepitemur e
10	24.5	43.8	9	2 Q94NA9	Q94NA9 daubentonla
11	24.5	43.8	9	2 Q94NB0	Q94NB0 microcebus
12	24.5	43.8	9	2 Q94NB1	Q94NB1 microcebus
13	24.5	43.8	9	2 Q94NB2	Q94NB2 microcebus
14	24.5	43.8	9	2 Q94XE6	Q94XE6 tectocoris
15	24	42.9	10	2 Q8SHN1	Q8SHN1 bradypodion
16	23	41.1	10	2 Q8SHC6	Q8SHC6 furcifer be
17	22	39.3	10	2 Q79912	Q79912 chamaeleo f
18	22	39.3	10	2 Q76ML6	Q76ML6 eurypharynx
19	22	39.3	10	2 Q9G697	Q9G697 chamaeleo d
20	22	39.3	10	2 Q8SH83	Q8SH83 brookesia t
21	22	39.3	10	2 Q8SH85	Q8SH85 brookesia t
22	22	39.3	10	2 Q8SH88	Q8SH88 brookesia t
23	22	39.3	10	2 Q8SH90	Q8SH90 brookesia s
24	22	39.3	10	2 Q8SH96	Q8SH96 brookesia p
25	22	39.3	10	2 Q8SHA2	Q8SHA2 brookesia b
26	22	39.3	10	2 Q8SHA5	Q8SHA5 brookesia a
27	22	39.3	10	2 Q8SHC9	Q8SHC9 furcifer ba
28	22	39.3	10	2 Q8SHD2	Q8SHD2 chamaeleo w
29	22	39.3	10	2 Q8SHD5	Q8SHD5 chamaeleo s
30	22	39.3	10	2 Q8SHD8	Q8SHD8 chamaeleo r
31	22	39.3	10	2 Q8SHE1	Q8SHE1 chamaeleo q

32	22	39.3	10	2 Q8SHE4	Q8SHE4 chamaeleo q
33	22	39.3	10	2 Q8SHE7	Q8SHE7 chamaeleo p
34	22	39.3	10	2 Q8SHF3	Q8SHF3 chamaeleo m
35	22	39.3	10	2 Q8SHF9	Q8SHF9 chamaeleo j
36	22	39.3	10	2 Q8SHG5	Q8SHG5 chamaeleo h
37	22	39.3	10	2 Q8SHG8	Q8SHG8 chamaeleo g
38	22	39.3	10	2 Q8SHH1	Q8SHH1 chamaeleo f
39	22	39.3	10	2 Q8SHH4	Q8SHH4 chamaeleo e
40	22	39.3	10	2 Q8SHH7	Q8SHH7 chamaeleo d
41	22	39.3	10	2 Q8SHI0	Q8SHI0 chamaeleo c
42	22	39.3	10	2 Q8SHI3	Q8SHI3 chamaeleo c
43	22	39.3	10	2 Q8SHI6	Q8SHI6 chamaeleo c
44	22	39.3	10	2 Q8SHI9	Q8SHI9 chamaeleo c
45	22	39.3	10	2 Q8SHJ2	Q8SHJ2 chamaeleo a

ALIGNMENTS

RESULT 1
LABA_JATMU
ID LABA_JATMU STANDARD; PRT; 10 AA.
AC P13270;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Labaditin.
OS Jatropa multifida (Physic nut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Malpighiales; Euphorbiaceae; Crotonoideae; Jatrophaeae;
OC Jatropa.
OX NCBI_TaxID=3996;
RN [1]
RP SEQUENCE.
RC TISSUE=Latex;
RA Kossai S., van der Sluis W.G., Boelens R., T'Hart L.A., Labadie R.P.;
RT "Labaditin, a novel cyclic decapeptide from the latex of Jatropa
multifida L. (Euphorbiaceae). Isolation and sequence determination by
RT means of two-dimensional NMR."
RL FEBS Lett. 256:91-96(1989).
CC -!- FUNCTION: Labaditin is an active peptide which inhibits the
CC classical pathway of complement activation in vitro. Activity
CC seems to be based on an interaction with C1.
CC -!- PTM: This is a cyclic peptide.
CC -!- MISCELLANEOUS: Latex of this plant is used in folkloric medicine
CC for treatment of infected wounds, skins infections and scabies.
KW Direct protein sequencing.
SQ SEQUENCE 10 AA; 1089 MW; D98AAD6362D1B362 CRC64;

Query Match 58.0%; Score 32.5; DB 1; Length 10;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 XGYWLTWGX 10
:|:|:|:
Db 1 AGVW-TVWGT 9

RESULT 2
Q70Y57 PRELIMINARY; PRT; 8 AA.
ID Q70Y57;
AC Q70Y57; (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Ribosomal protein (fragment).
GN Name-rps16;
OS Fuerstia africana.

OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Lamiales; Lamiaceae; Nepetoideae; Ocimeae; Fuerstia.

```

OC NCBI_TaxID=204226;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15019625;
RA Paton A., Springate D.A., Sudde S., Otieno D., Grayer R., Harley M.M.,
RA Willis F., Simonds M.S.J., Powell M.P., Savolainen V.;
RT "Phylogeny and evolution of baobabs and allies (Ocimeae, Labiatae)
RT based on three plastid DNA regions.";
RL Mol. Phylogenet. Evol. 31:277-299(2004).
DR EMBL; AJ505427; CAD45547.1; -.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW Ribosomal protein.
FT NON_TER 1
FT NON_TER 8
FT NON_TER 8
SQ SEQUENCE 8 AA; 916 MW; DABEAB58637041B5 CRC64;

Query Match 48.2%; Score 27; DB 2; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIWGX 10
DB 2 TIWGY 6

RESULT 3
CAD45547
ID CAD45547 PRELIMINARY; PRT; 8 AA.
AC CAD45547;
DT 14-MAR-2004 (TReMBLrel. 27, Created)
DT 14-MAR-2004 (TReMBLrel. 27, Last sequence update)
DE Ribosomal protein (Fragment).
GN Rps16.
OS Ficus africana.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiaceae; Nepetoideae; Ocimeae; Fuerstia.
OX NCBI_TaxID=204226;
RN [1]
RP SEQUENCE FROM N.A.
RA Paton A., Springate D.A., Sudde S., Otieno D., Grayer R., Harley M.M.,
RA Willis F., Simonds M.S.J., Powell M.P., Savolainen V.;
RT "Phylogeny and evolution of baobabs and allies (Ocimeae, Labiatae)
RT based on three plastid DNA regions.";
RL Mol. Phylogenet. Evol. 31:277-299(2004).
DR EMBL; AJ505427; CAD45547.1; -.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW Ribosomal protein.
FT NON_TER 1
FT NON_TER 8
FT NON_TER 8
SQ SEQUENCE 8 AA; 916 MW; DABEAB58637041B5 CRC64;

Query Match 48.2%; Score 27; DB 2; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIWGX 10
DB 2 TIWGY 6

RESULT 4
QSHF6
ID QSHF6 PRELIMINARY; PRT; 10 AA.
AC QSHF6;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN Name=COI;
OS Chamaeleo melleri (Meller's chameleon).
OG Mitochondrion.
```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae; Chamaeleo.
OX NCBI_TaxID=179915;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22169767; PubMed=12182400;
RA Townsend T., Larson A.;
RT "Molecular phylogenetics and mitochondrial genomic evolution in the
RT chamaeleonidae (Reptilia, Squamata).";
RL Mol. Phylogenet. Evol. 23:22-36(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Townsend T.M., Larson A.L.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF448755; AAL90547.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 10
FT NON_TER 10
SQ SEQUENCE 10 AA; 1276 MW; 5E218E2733772727 CRC64;

Query Match 48.2%; Score 27; DB 2; Length 10;
Best Local Similarity 60.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WLTIW 8
DB 2 WLLIW 6

RESULT 5
Q7M3E8
ID Q7M3E8 PRELIMINARY; PRT; 10 AA.
AC Q7M3E8;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Glutathione transferase (EC 2.5.1.18) class mu 4 (Fragment).
OS Sus scrofa domestica (domestic pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9825;
RN [1]
RP SEQUENCE.
RA Rouimi P., Anglade P., Debrauwer L., Tulliez J.;
RT "Characterization of pig liver glutathione S-transferases using HPLC-
RT electrospray-ionization mass spectrometry.";
RL Biochem. J. 317:879-884 (1996).
DR PIR; S71868; S71868.
DR GO; GO:0004364; F:glutathione transferase activity; IEA.
FT NON_TER 10
FT NON_TER 10
SQ SEQUENCE 10 AA; 1223 MW; 5E16395AB36B5877 CRC64;

Query Match 44.6%; Score 25; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYW 4
DB 3 LGYW 6

RESULT 6
Q6JL97
ID Q6JL97 PRELIMINARY; PRT; 10 AA.
AC Q6JL97;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Nuol (Fragment).
GN Name=nuol;
OS Neisseria gonorrhoeae.
OS Neisseria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
```



```

OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MS11;
RX PubMed=15084227;
RA Snyder L.A., Davies J.K., Saunders N.J.;
RT "Microarray genotyping of key experimental strains of Neisseria
RT gonorrhoeae reveals gene complement diversity and five new neisserial
RT genes associated with Minimal Mobile Elements.";
RL BMC Genomics 5:23-23(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MS11;
RA Snyder L.A.S., Davies J.K., Saunders N.J.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY386266; AAS16521.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1227 MW; BACCB286379D1A6 CRC64;

Query Match 44.6%; Score 25; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 LTIWGX 10
DB 3 MTFWGL 8

RESULT 7
AAS16521
ID AAS16521 PRELIMINARY; PRT; 10 AA.
AC AAS16521;
DT 20-MAY-2004 (TrEMBLrel. 27, Created)
DT 20-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 20-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE NuOL.
GN Neisseria gonorrhoeae.
OS Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MS11;
RX PubMed=15084227;
RA Snyder L.A., Davies J.K., Saunders N.J.;
RT "Microarray genotyping of key experimental strains of Neisseria
RT gonorrhoeae reveals gene complement diversity and five new neisserial
RT genes associated with Minimal Mobile Elements.";
RL BMC Genomics 5:23-23(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MS11;
RA Snyder L.A.S., Davies J.K., Saunders N.J.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY386266; AAS16521.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1227 MW; BACCB286379D1A6 CRC64;

Query Match 44.6%; Score 25; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 LTIWGX 10
DB 3 MTFWGL 8

RESULT 8
Q85DB0
ID Q85DB0 PRELIMINARY; PRT; 9 AA.
AC Q85DB0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

```

```

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome oxidase subunit III (Fragment).
GN Name=COIII;
OS Lepilemur septentrionalis (northern sportive lemur).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Megaladapidae; Lepilemur.
OX NCBI_TaxID=78584;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22631663; PubMed=12719521;
RA Pastorini J., Thalmann U., Martin R.D.;
RT "A molecular approach to comparative phylogeography of extant Malagasy
RT lemurs.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5879-5884(2003).
DR EMBL: AF224597; AAP33652.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1174 MW; 16C563636B5045B0 CRC64;

Query Match 43.8%; Score 24.5; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

QY 3 YWLTWGX 10
DB 5 YW---WGS 9

RESULT 9
Q85DB8
ID Q85DB8 PRELIMINARY; PRT; 9 AA.
AC Q85DB8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome oxidase subunit III (Fragment).
GN Name=COIII;
OS Lepilemur edwardsi (Milne-Edwards's sportive lemur).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Megaladapidae; Lepilemur.
OX NCBI_TaxID=122230;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22631663; PubMed=12719521;
RA Pastorini J., Thalmann U., Martin R.D.;
RT "A molecular approach to comparative phylogeography of extant Malagasy
RT lemurs.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5879-5884(2003).
DR EMBL: AF224595; AAP33644.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 43.8%; Score 24.5; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

QY 3 YWLTWGX 10
DB 5 YW---WGS 9

RESULT 10
Q94NA9
ID Q94NA9 PRELIMINARY; PRT; 9 AA.
AC Q94NA9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

```


GN Name=COIII;
OS Microcebus murinus (Lesser mouse lemur).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirrhini; Cheirogaleidae;
OC Microcebus.
OX NCBI_TaxID=30608;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=21184272; PubMed=11286490;
RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
RT "Molecular phylogeny of the lemur family cheirogaleidae (primates)
RT based on mitochondrial DNA sequences."
RL Mol. Phylogenet. Evol. 19:45-56(2001).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=22631663; PubMed=12719521;
RA Pastorini J., Thalmann U., Martin R.D.;
RT "A molecular approach to comparative phylogeography of extant Malagasy
RT lemurs."
RL Proc. Natl. Acad. Sci. U.S.A. 100:5879-5884(2003).
DR EMBL; AF224624; AAK70547.1; -;
DR EMBL; AF224625; AAK70551.1; -;
DR EMBL; AF224626; AAK70555.1; -;
DR EMBL; AF224627; AAK70559.1; -;
DR EMBL; AF224628; AAK70563.1; -;
DR EMBL; AF224629; AAK70567.1; -;
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;
Query Match 43.8%; Score 24.5; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 3; Gaps 1;
OY 3 YWLTWGX 10
DB 5 YW---WGS 9
RESULT 14
OY Q94XE6 PRELIMINARY; PRT; 9 AA.
AC Q94XE6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit III (Fragment).
GN Name=cx3;
OS Tectocoris diopthalmus (cotton harlequin bug).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC Panheteroptera; Pentatomomorpha; Pentatomoidea; Pentatomidae;
OC Tectocoris.
OX NCBI_TaxID=159956;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=21396409; PubMed=11504862;
RA Shao R., Campbell N.J., Schmidt E.R., Barker S.C.;
RT "Increased rate of gene rearrangement in the mitochondrial genomes of
RT three orders of hemipteroid insects."
RL Mol. Biol. Evol. 18:1828-1832(2001).
DR EMBL; AF335990; AAK55283.1; -;
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1206 MW; A2C563636B5041A6 CRC64;
Query Match 43.8%; Score 24.5; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

OY 3 YWLTWGX 10
DB 5 YW---WGS 9
RESULT 15
OY Q8SHN1 PRELIMINARY; PRT; 10 AA.
AC Q8SHN1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN Name=COI;
OS Bradypodion tavetanum (Dwarf Fischer's chameleon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylia; Squamata; Iguania; Acrodonta; Chamaeleonidae;
OC Bradypodion.
OX NCBI_TaxID=179888;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=22169767; PubMed=12182400;
RA Townsend T., Larson A.;
RT "Molecular phylogenetics and mitochondrial genomic evolution in the
RT chamaeleonidae (Reptilia, Squamata)."
RL Mol. Phylogenet. Evol. 23:22-36(2002).
RN (2)
RP SEQUENCE FROM N.A.
RA Townsend T.M., Larson A.L.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF448730; AAL90472.1; -;
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1327 MW; 5E2180C7336415B7 CRC64;
Query Match 42.9%; Score 24; DB 2; Length 10;
Best Local Similarity 60.0%; Pred. No. 4.5e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 4 WLTW 8
DB 2 WLSRW 6
Search completed: January 3, 2005, 17:00:41
Job time : 185 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 12:55:07 ; Search time 69.6226 Seconds
(without alignments)
51.525 Million cell updates/sec

Title: US-10-046-922-35

Perfect score: 72

Sequence: 1 CGYWLTTWGC 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	72	100.0	10	5	ABP53932	Abp53932	VEGFR-3 b
2	54	75.0	10	5	ABP53931	Abp53931	VEGFR-3 b
3	50	69.4	10	5	ABP53968	Abp53968	VEGFR-3 b
4	47.5	66.0	304	7	ADC87481	Adc87481	Human GPC
5	47	65.3	25	7	ADC99638	Adc99638	Cancer-re
6	46	63.9	17	6	AAO26074	Aao26074	Fc region
7	46	63.9	17	8	ADJ50741	Adj50741	Human ser
8	46	63.9	136	8	ADM87650	Adm87650	Human EST
9	45	62.5	20	5	AAU90545	Aau90545	Insulin/i
10	45	62.5	129	2	AAV59880	Aav59880	Human nor
11	45	62.5	828	6	ABU49938	Abu49938	Protein e
12	44.5	61.8	1024	5	ABBO4861	Abbo4861	LDL recep
13	44	61.1	149	8	ADP29850	Adp29850	Human sec
14	44	61.1	166	6	AAE31487	Aae31487	Human but
15	44	61.1	415	5	AAU91293	Aau91293	Human NOV
16	43	59.7	13	6	AAO26093	Aao26093	Fc region
17	43	59.7	13	8	ADJ50760	Adj50760	Human ser
18	43	59.7	454	5	ABP73979	Abp73979	Candida a
19	43	59.7	474	6	ABU30004	Abu30004	Protein e
20	43	59.7	492	7	ADC97318	Adc97318	E. faeciu
21	43	59.7	1084	8	ADL1869	Adl1869	P. aerugi
22	42.5	59.0	152	3	AAV94989	Aav94989	Human sec
23	42.5	59.0	152	5	ABB90081	Abb90081	Human pol
24	42.5	59.0	152	5	AAO17173	Aao17173	Human sec
25	42.5	59.0	152	5	ABG64784	Abg64784	Human alb

26	42.5	59.0	152	8	ADL78051	Adl78051	Albumin f
27	42.5	59.0	159	2	AAE66278	Aae66278	Therapob
28	42.5	59.0	159	3	AAE12156	Aae12156	Hydrophob
29	42.5	59.0	159	4	AAE78581	Aae78581	Human pro
30	42.5	59.0	159	4	AAE73100	Aae73100	Human ang
31	42.5	59.0	159	5	AAO17198	Aao17198	Human sec
32	42.5	59.0	159	5	ABG64785	Abg64785	Human alb
33	42.5	59.0	159	8	ADL78052	Adl78052	Albumin f
34	42.5	59.0	159	8	ADN05188	Adn05188	Antipeori
35	42.5	59.0	161	4	AAE25822	Aae25822	Human pro
36	42.5	59.0	161	4	ABE12006	Abe12006	Human gli
37	42.5	59.0	161	4	AAE79565	Aae79565	Human pro
38	42.5	59.0	176	6	ABO07116	Ab007116	Novel hum
39	42.5	59.0	1024	5	ABBO4863	Abbo4863	LDL recep
40	42.5	59.0	1765	2	AAV16572	Aav16572	Type 5 so
41	42.5	59.0	1765	2	AAV41668	Aav41668	Rat senso
42	42.5	59.0	1765	2	AAV06596	Aav06596	Rat sodiu
43	42.5	59.0	1765	4	AAE20122	Aae20122	Rat sodiu
44	42.5	59.0	1765	4	AAE20123	Aae20123	Rat sodiu
45	42.5	59.0	1765	7	ADD32192	Add32192	Rat Na v

ALIGNMENTS

RESULT 1

ABP53932
ID ABP53932 standard; peptide; 10 AA.

AC ABP53932;

DT 09-JAN-2003 (first entry)

DE VEGFR-3 binding peptide SEQ ID NO:35.

XX Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
KW cytosolic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
KW vulnary; cell surface receptor; cancer; neovascularisation;
KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
KW diabetes; PDGF; platelet derived growth factor.

OS Homo sapiens.

OS Synthetic.

PN WO200257299-A2.

XX 25-JUL-2002.

PD 16-JAN-2002; 2002WO-IB0000099.

XX 17-JAN-2001; 2001US-0262476P.

XX (LUDW-) LUDWIG INST CANCER RES.

XX (LICN) LICENTIA LTD.

XX Alitalo K, Koivunen E, Kubo H;

XX WPI; 2002-691521/74.

XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
such as cancer and diseases of neovascularization.

XX Claim 13; Page 80; 149pp; English.

XX The present invention describes an isolated peptide (I) that binds to and
inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
CC have cytostatic, hepatotropic, antitumor, antineoplastic, hypotensive,
CC antidiabetic and vulnary activities, and can be used in gene therapy.
CC Compositions and methods from the present invention are useful for
diagnosing, evaluating and treating disorders mediated by the activity of
the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 XX
 SQ Sequence 10 AA;

Query Match 100.0%; Score 72; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGYWLTIWGC 10
 |||||
 Db 1 CGYWLTIWGC 10

RESULT 2
 ABP53931
 ID ABP53931 standard; peptide; 10 AA.

XX AC ABP53931;

DT 09-JAN-2003 (first entry)

XX DE VEGFR-3 binding peptide SEQ ID NO:34.

XX KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
 KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
 KW cytosolic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
 KW vulnary; cell surface receptor; cancer; neovascularisation;
 KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
 KW diabetes; PDGF; platelet derived growth factor.

XX OS Homo sapiens.
 OS Synthetic.

XX FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "any amino acid"
 FT Misc-difference 10 /note= "any amino acid"

XX WO200257299-A2.

XX 25-JUL-2002.

XX PF 16-JAN-2002; 2002WO-IB000099.

XX PR 17-JAN-2001; 2001US-0262476P.

XX PA (LUDW-) LUDWIG INST CANCER RES.
 XX PA (LICN) LICENTIA LTD.

XX PI Alitalo K, Koivunen E, Kubo H;

XX WPI; 2002-691521/74.

XX PT New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
 PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
 PT such as cancer and diseases of neovascularization.

XX PS Claim 12; Page 80; 149pp; English.

XX CC The present invention describes an isolated peptide (I) that binds to and
 CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnary activities, and can be used in gene therapy.
 CC Compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 XX
 SQ Sequence 10 AA;

Query Match 75.0%; Score 54; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.67;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GYWLTIWG 9
 |||||
 Db 2 GYWLTIWG 9

RESULT 3
 ABP53968
 ID ABP53968 standard; peptide; 10 AA.

XX AC ABP53968;

DT 09-JAN-2003 (first entry)

XX DE VEGFR-3 binding peptide SEQ ID NO:73.

XX KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
 KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
 KW cytosolic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
 KW vulnary; cell surface receptor; cancer; neovascularisation;
 KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
 KW diabetes; PDGF; platelet derived growth factor.

XX OS Homo sapiens.
 OS Synthetic.

XX FH Key Location/Qualifiers
 FT Misc-difference 5.7 /note= "X is any amino acid"
 FT Misc-difference 9 /note= "X is any amino acid"

XX WO200257299-A2.

XX 25-JUL-2002.

XX PF 16-JAN-2002; 2002WO-IB000099.

XX PR 17-JAN-2001; 2001US-0262476P.

XX PA (LUDW-) LUDWIG INST CANCER RES.
 XX PA (LICN) LICENTIA LTD.

XX PI Alitalo K, Koivunen E, Kubo H;

XX WPI; 2002-691521/74.

XX PT New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
 PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
 PT such as cancer and diseases of neovascularization.

XX PS Disclosure; Page 147; 149pp; English.

XX CC The present invention describes an isolated peptide (I) that binds to and
 CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnary activities, and can be used in gene therapy.
 CC Compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a VEGFR-3 binding peptide, which is given in the
 CC exemplification of the present invention

XX Sequence 10 AA;

Query Match 69.4%; Score 50; DB 5; Length 10;
 Best Local Similarity 60.0%; Pred. No. 2.3;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGYWLTIWGC 10
 |||||
 Db 1 CGYWXWXC 10

RESULT 4

ADC87481
 ID ADC87481 standard; protein; 304 AA.

XX AC ADC87481;

XX DT 01-JAN-2004 (first entry)

XX DE Human GPCR protein SEQ ID NO:1934.

XX KW human; GPCR; guanosine triphosphate-binding protein coupled receptor;
 KW gene therapy.

XX OS Homo sapiens.

XX PN EP1270724-A2.

XX PD 02-JAN-2003.

XX PF 18-JUN-2002; 2002EP-00013517.

XX PR 18-JUN-2001; 2001JP-00246789.

XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

XX PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.

XX PI Suwa M, Asai K, Akiyama Y, Aburatani H;

XX DR WPI; 2003-315783/31.

XX DR N-PSDB; ADC87480.

XX PT New polynucleotide, useful for preparing a composition for treating a
 PT patient in need of increased or suppressed activity or expression of the
 PT guanosine triphosphate-binding protein coupled receptor.

PS Claim 2; SEQ ID NO 1934; 28pp; English.

XX CC The invention relates to a novel polynucleotide encoding a guanosine
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
 CC the invention may have a use in gene therapy. The polynucleotide and
 CC polypeptide are useful for preparing a composition for treating a patient
 CC in need of increased or suppressed activity or expression of the
 CC guanosine triphosphate-binding protein coupled receptor. The protein
 CC sequences shown in ADC85549-ADC87617 represent GPCR's of the invention.

XX Sequence 304 AA;

Query Match 66.0%; Score 47.5; DB 7; Length 304;

Best Local Similarity 60.0%; Pred. No. 91;

Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 CGYWLTIWGC 10
 |||||
 Db 108 CGFW-AVWGC 116

RESULT 5

ADC99638

ID ADC99638 standard; peptide; 25 AA.

XX AC ADC99638;

XX DT 01-JAN-2004 (first entry)

XX DE Cancer-related Tie-1-binder peptide - SEQ ID 476.

XX KW cytostatic; cancer; gene therapy; DGI-2; DGI-5; DGI-7; DGI-9; Hras;
 KW lepton; VEGF; vascular endothelial growth factor receptor; VEGF-R1;
 KW VEGF-R2; VEGF-R3; FLT1; FMS-related tyrosine kinase 1; FLK1; KDR;
 KW kinase insert domain protein receptor; EGFR; epidermal growth factor;
 KW FGFR1; fibroblast growth factor; Tie-1.

XX OS Unidentified.

XX PN WO2003035839-A2.

XX PD 01-MAY-2003.

XX PF 24-OCT-2002; 2002WO-US034021.

XX PR 24-OCT-2001; 2001US-0345471P.

XX PA (DGI-B-) DGI BIOTECHNOLOGIES INC.

XX PI Pillutia RC, Brissette R, Spruyt M, Dedova O, Blume A;

XX PI Prendergast J, Goldstein N;

XX DR WPI; 2003-457332/43.

XX PT Selecting target and target binder pairs for preparing a composition for
 PT treating cancer by mixing in a reaction vessel phase expressing
 PT biological targets and phase expressing target binders.

XX PS Claim 26; SEQ ID NO 476; 172pp; English.

XX CC The invention relates to a novel method of selecting target and target
 CC binder pairs comprising mixing in a reaction vessel phase expressing
 CC biological targets and phase expressing target binders, each having
 CC distinguishable selection markers and selecting target and target binder
 CC pairs based on the selection markers. The molecules of the invention
 CC demonstrate cytostatic activity whilst the method may be useful for
 CC selecting target and target binder pairs for preparing a composition for
 CC treating cancer. Furthermore, the method may be utilised during gene
 CC therapy procedures. The current sequence is that of the cancer-related
 CC Tie1-binder peptide of the invention.

XX Sequence 25 AA;

Query Match 65.3%; Score 47; DB 7; Length 25;

Best Local Similarity 66.7%; Pred. No. 13;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGYWLTIWGC 9

Db 5 CGYWGELMG 13

RESULT 6

AAO26074

ID AAO26074 standard; peptide; 17 AA.

XX AC AAO26074;

XX DT 03-APR-2003 (first entry)

XX DE Fc region binding peptide SEQ ID No 54.

XX

KW Immunoglobulin Fc region; binding; whole blood; plasma; transgenic milk;
KW antibody response; half-life; stability; circulatory system.
XX Unidentified.
OS WO200286070-A2.
XX 31-OCT-2002.
XX 18-APR-2002; 2002WO-US012492.
XX 18-APR-2001; 2001US-0284534P.
XX (DYAX-) DYAX CORP.
XX Rondon IU, Wu Q, Ley AC, Stochl M, Ransohoff TC, Potter MD;
XX WPI; 2003-201220/19.
XX New polypeptides, useful as binding molecules for detecting, isolating or
PT purifying immunoglobulin Fc-region polypeptides present in a solution, or
PT for regulating or preventing an antibody response.
XX Claim 3; Page 76; 152pp; English.
XX The invention relates to novel isolated polypeptides comprising a
CC sequence that binds an immunoglobulin Fc region. The polypeptides are
CC useful as binding molecules for detecting, isolating or purifying
CC immunoglobulin Fc-region polypeptides present in a solution, e.g. whole
CC blood, plasma or transgenic milk. The Fc-region binding polypeptides are
CC also useful for regulating or preventing an antibody response or for
CC increasing the half-life and over all stability of a therapeutic or
CC diagnostic compound that is administered to or enters the circulatory
CC system of an individual. This sequence represents an Fc region binding
CC peptide of the invention
XX
XX Sequence 17 AA;
SQ
Query Match 63.9%; Score 46; DB 6; Length 17;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGYWLTIWG 9
Db ||:|||||
4 CGFWPRIWG 12
RESULT 7
ADJ50741
ID ADJ50741 standard; peptide; 17 AA.
XX
AC ADJ50741;
XX
XX 06-MAY-2004 (first entry)
XX
XX Human serum albumin binding peptide, Seq ID No 278.
DE human serum albumin; HSA; serum; blood; tumour; human.
XX
XX Homo sapiens.
OS
XX WO2003106493-A1.
XX
XX 24-DEC-2003.
XX
XX 16-JUN-2003; 2003WO-US018896.
XX
XX 14-JUN-2002; 2002US-0388642P.
XX
XX (DYAX-) DYAX CORP.
XX
XX Sato AK, Dawson BM;
XX

DR WPI; 2004-082161/08.
XX
PT Evaluating sample comprising soluble serum protein by forming complex
PT comprising serum protein and physically associated compounds using
PT peptide ligand that specifically binds with proteins, which is separated
PT and evaluated.
XX
XX Disclosure; SEQ ID NO 278; 191pp; English.
XX
CC The invention relates to a method of evaluating sample by providing a
CC soluble serum protein (I), one or more compounds physically associated
CC with (I), and a (I)-binding agent that comprises a peptide that
CC specifically binds to (I), allowing the (I)-binding agent to bind to (I)
CC to form a complex including one or more compounds physically associated
CC with (I), separating the complex from one or more components of the
CC sample, and evaluating one or more of the physically associated
CC compounds. The sample comprises blood or serum, or is obtained from a
CC biopsy. The sample may also be obtained from a tumour or a region within
CC 5 mm of a tumour. The method is useful for detecting modulators that
CC modulate interaction of serum protein-binding compound and serum protein
CC and for identifying binding ligands for serum protein. The present
CC sequence represents a serum albumin-binding peptide identified using the
CC method of the invention.
XX
XX Sequence 17 AA;
SQ
Query Match 63.9%; Score 46; DB 8; Length 17;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGYWLTIWG 9
Db ||:|||||
4 CGFWPRIWG 12
RESULT 8
ADM87650
ID ADM87650 standard; protein; 136 AA.
XX
AC ADM87650;
XX
XX 03-JUN-2004 (first entry)
XX
XX Human EST derived amino acid sequence SEQ ID NO:743.
DE
XX
XX respiratory; cytostatic; antiarthritic; antiinflammatory;
KW gastrointestinal; antibacterial; immunosuppressive; antidiabetic;
KW antirheumatic; gene therapy; molecular weight marker; chromosome marker;
KW chromosome tag; genetic fingerprinting; nutritional supplement; cancer;
KW inflammatory condition; arthritis; inflammatory bowel disease;
KW Crohn's disease; sepsis; rheumatoid arthritis; diabetes mellitus type 1;
KW graft versus host disease; human; expressed sequence tag; EST.
XX
XX Homo sapiens.
OS
XX WO2004009834-A2.
XX
XX 29-JAN-2004.
XX
XX 19-JUL-2002; 2002WO-US022858.
XX
XX 21-JUL-2001; 2001US-0306971P.
XX
XX 28-MAR-2002; 2002US-00112944.
XX
XX (NUVE-) NUVELO INC.
XX
XX Tang YT, Yang Y, Weng G, Zhang J, Ren F, Xue A, Wang J;
PI Wehrman T, Ghosh MJ, Wang D, Zhao QA, Wang Z;
XX
XX WPI; 2004-143291/14.
DR N-PSDB; ADM87432.
XX
XX New isolated polynucleotides and polypeptides, useful for treating, e.g.

PT cancer, lung or liver fibrosis, arthritis, inflammatory bowel disease,
 PT Crohn's disease, rheumatoid arthritis, diabetes mellitus type 1 or graft
 PT versus host disease.

PS Example 2; SEQ ID NO 743; 591pp; English.

XX The present invention describes an isolated polynucleotide (I): (a)
 CC comprising a nucleotide sequence selected from SEQ ID NO:1-244; or (b)
 CC which encodes a polypeptide with biological activity, where the
 CC polynucleotide hybridises to (I) under stringent hybridisation conditions
 CC or has greater than 99% sequence identity with (I). (I) has respiratory,
 CC cytostatic, antiarthritic, antiinflammatory, gastrointestinal,
 CC antibacterial, immunosuppressive, antidiabetic and antirheumatic
 CC activities, and can be used in gene therapy. (I) can be used for
 CC generating polynucleotides encoding chimeric or fusion proteins and
 CC heterologous protein sequences. The polynucleotides can be used to
 CC express recombinant protein for analysis, characterisation or therapeutic
 CC use; as markers for tissues in which the corresponding protein is
 CC preferentially expressed; as molecular weight markers on gels; as
 CC chromosome markers or tags to identify chromosomes or to map related gene
 CC positions; to compare with endogenous DNA sequences in patients to
 CC identify potential genetic disorders; as probes to hybridise and discover
 CC genes, related DNA sequences; as a source of information to derive PCR
 CC primers for genetic fingerprinting; as a probe to subtract-out known
 CC sequences in the process of discovering other novel polynucleotides; for
 CC selecting and making oligomers for attachment to a gene chip or other
 CC support, including for examination of expression patterns; to raise anti-
 CC protein antibodies using DNA immunisation techniques; and as an antigen
 CC to raise anti-DNA antibodies or elicit another immune response. The
 CC polynucleotides and polypeptides can also be used as nutritional sources
 CC or supplements, e.g. as a protein or amino acid supplement, as a carbon
 CC source, as a nitrogen source or as a source of carbohydrates. The
 CC polynucleotides and polypeptides can also be used to treat cancer. The
 CC compositions are useful for promoting better or faster closure of non-
 CC healing wounds, for the generation and regeneration of tissues, for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, and conditions resulting from
 CC systemic cytokine damage. The compositions can also be used to treat
 CC inflammatory conditions (e.g. arthritis, inflammatory bowel disease or
 CC Crohn's disease), sepsis, rheumatoid arthritis, diabetes mellitus type 1
 CC or graft versus host disease. The present sequence represents an
 CC expressed sequence tag (EST) derived amino acid sequence from the present
 CC invention. N.B. The sequences for this patent were obtained from the
 CC USPTO web site from an equivalent US patent US20040048249A1.

XX Sequence 136 AA;

Query Match 63.9%; Score 46; DB 8; Length 136;
 Best Local Similarity 60.0%; Pred. No. 73;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 CGYWLITWGC 10
 ||||| :|||
 Db 93 CGRDWLWGC 102

RESULT 9

AAU90545
 ID AAU90545 standard; peptide; 20 AA.

XX AC AAU90545;

XX 18-JUN-2002 (first entry)

DE Insulin/insulin-like growth factor receptor-binding peptide #2501.

XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;

KW ophthalmological; insulin; receptor; gene therapy; diabetes;

KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;

KW diabetic retinopathy; neurological diseases; stroke; diabetic neuropathy.

XX Synthetic.

XX

PN WO200172771-A2.

XX 04-OCT-2001.

XX 29-MAR-2000; 2000WO-US008528.

XX 29-MAR-2000; 2000WO-US008528.

PR (DGTB-) DGI BIOTECHNOLOGIES LLC.

PA (NOVO) NOVO NORDISK AS.

XX Beasley J, Blume AJ, Schaeffer L, Pillutla R, Brandt J;

PI Briesette R, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;

PI Hansen PH, Ravera M, Hsiao K;

XX WPI; 2002-025774/03.

XX Modulating insulin activity in mammalian cells, for treating e.g.
 PT diabetes and tumors, comprises using peptides that bind to insulin or
 PT insulin-like growth factor receptors.

PS Disclosure; Fig 8-3; 390pp; English.

XX The invention relates to a method of modulating insulin activity in
 CC mammalian cells by administering a peptide that binds the insulin
 CC receptor (IR). A composition containing a peptide, optionally expressed
 CC from gene therapy vectors, that binds to Site 1 of IR and an insulin
 CC agonist are useful for treating diabetes. Also, peptides that are
 CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are
 CC useful for treating insulin-like growth factor (IGF)-sensitive tumours
 CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1
 CC receptor agonists are useful for treating neurological diseases,
 CC including stroke and diabetic neuropathy. The peptides are also useful in
 CC screening for compounds that bind to IR or IGF-1 receptor, potential
 CC therapeutics and research reagents. AAU8034-AAU90957 represent IR and/or
 CC IGF-1 receptor-binding peptides and related amino acid sequences of the
 CC invention

XX Sequence 20 AA;

Query Match 62.5%; Score 45; DB 5; Length 20;
 Best Local Similarity 60.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CGYWLITWGC 10
 ||||| :|||
 Db 8 CGAWPTTWNC 17

RESULT 10

AAU59880
 ID AAU59880 standard; protein; 129 AA.

XX AC AAU59880;

XX 19-JAN-2000 (first entry)

DE Human normal uterus tissue derived protein 43.

XX Human; uterus; cancer; treatment; anticancer; cytostatic; gene therapy;

KW EST; expressed sequence tag.

XX Homo sapiens.

XX DE19817946-A1.

XX 21-OCT-1999.

XX 17-APR-1998; 98DE-01017946.

XX 17-APR-1998; 98DE-01017946.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

```
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl B;
XX WPI; 1999-591956/51.
XX N-PSDB; AAZ41339.
XX
XX New nucleic acid sequences expressed in normal uterine tissues, and
XX derived polypeptides, for treatment of uterine cancer and identification
XX of therapeutic agents.
XX
XX Claim 23; Page 138; 154pp; German.
XX
XX This invention describes novel cDNA sequences (A) highly expressed in
XX normal uterine tissue which can have anticancer and cytostatic activity
XX and can be used for gene therapy. (A) are used (i) for recombinant
XX expression of polypeptides (B) and (ii) to isolate complete genes. (B)
XX are used (i) to identify agents suitable for treatment of uterine cancer;
XX (ii) directly for treating this form of cancer (including expression from
XX gene therapy vectors) and (iii) for generation of specific antibodies.
XX (A) are identified by assembling ESTs (expressed sequence tags) from a
XX particular tissue type before comparison of expression patterns. This
XX allows a significantly longer fragment of the gene to be revealed, so
XX should reduce the number of failures associated with the fact that ESTs
XX from different libraries may represent different parts of the same
XX unknown gene, distorting the estimated frequency of occurrence in a
XX particular tissue. AAY59838-Y59892 represent protein fragments encoded by
XX the human uterine tissue derived cDNA fragments represented in AAZ41325-
XX Z41385
XX
XX Query Match 62.5%; Score 45; DB 2; Length 129;
XX Best Local Similarity 62.5%; Pred. No. 96;
XX Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 CGYWLTIW 8
XX | : ||| : |
XX 32 CSHWLTWV 39
XX
XX
XX RESULT 11
XX ABU49938
XX ID ABU49938 standard; protein; 828 AA.
XX
XX AC ABU49938;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by Prokaryotic essential gene #35465.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Yersinia pestis.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX
XX 06-SEP-2001; 2001US-00948993.
XX
XX 25-OCT-2001; 2001US-0342923P.
XX
XX 08-FEB-2002; 2002US-00072851.
XX
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX
XX N-PSDB; ACA53808.
XX
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 77862; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Query Match 62.5%; Score 45; DB 6; Length 828;
XX Best Local Similarity 58.3%; Pred. No. 4.6e+02;
XX Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
XX
XX 1 CGYWLTI--WGC 10
XX || || : |||
XX 372 CGLWLELLSWG 383
XX
XX
XX RESULT 12
XX ABB04861
XX ID ABB04861 standard; protein; 1024 AA.
XX
XX AC ABB04861;
XX
XX 13-MAR-2002 (first entry)
XX
XX LDL receptor binding protein Na channel brain 3 SEQ ID NO:84.
XX
XX Low density lipoprotein receptor binding protein; signal transduction;
XX LDL receptor binding protein; LDL receptor signalling pathway.
XX
XX Synthetic.
XX
XX WO2001184159-A2.
XX
XX 08-NOV-2001.
XX
XX 24-APR-2001; 2001WO-US013214.
XX
XX 01-MAY-2000; 2000US-00562737.
XX
```

PA (TEXA) UNIV TEXAS SYSTEM.
 XX Herz J, Gotthardt M;
 PI
 XX
 XX
 DR WPI; 2002-082855/11.
 XX
 PT Detecting stress that alters interaction of LDL receptor binding
 PT polypeptide with LDL receptor interaction domain, comprises detecting
 PT difference in stress-biased and unbiased interaction of peptide and
 PT domain in a system.
 XX
 XX
 PS Disclosure; Page 148-150; 200pp; English.
 XX
 CC The present invention describes a method for detecting a stress that
 CC alters a functional interaction of a low density lipoprotein (LDL)
 CC receptor binding protein (I) with an LDL receptor interaction domain
 CC (II). The method involves introducing a predetermined stress into a
 CC system which provides a stress-biased physical interaction of (I) with
 CC (II), where in the absence of the stress, the system provides an unbiased
 CC interaction of (I) and (II), and detecting the stress-biased interaction
 CC of (I) and (II), where a difference between BI and UI indicates that the
 CC stress alters the interaction of (I) and (II). (I) is selected from
 CC SEMCAP-1, JIP-1, PSD-95, JIP-2, Talin, OMP25, CAPON, PIP4,5 Kinase, Na
 CC channel brain 3, Mint1, ICAP-1 and APC subunit 10. The method is useful
 CC for detecting a stress that alters functional interaction of LDL receptor
 CC binding polypeptide with LDL receptor interaction domain. The method is
 CC useful for detecting and modulating signal transduction through LDL
 CC receptors. AB904778 to AB904909 represent LDL receptor binding proteins
 CC which are used in the exemplification of the present invention
 XX
 SQ Sequence 1024 AA;

Query Match 61.8%; Score 44.5; DB 5; Length 1024;
 Best Local Similarity 54.5%; Pred. No. 6.5e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 CGYWL-TWGC 10
 ||:|:|:|:
 Db 892 CGHWIETWDC 902

RESULT 13
 ADP29850
 ID ADP29850 standard; protein; 149 AA.
 XX
 AC ADP29850;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human secreted protein SEQ ID #617.
 XX
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
 KW cancer; inflammatory; immune; human secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO2004035732-A2.
 XX
 PD 29-APR-2004.
 XX
 PF 28-AUG-2003; 2003WO-US026780.
 XX
 PR 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406579P.
 PR 29-AUG-2002; 2002US-0406583P.
 PR 29-AUG-2002; 2002US-0406588P.
 PR 29-AUG-2002; 2002US-0406608P.
 PR 29-AUG-2002; 2002US-0406611P.
 PR 29-AUG-2002; 2002US-0406612P.
 PR 29-AUG-2002; 2002US-0406616P.
 PR 29-AUG-2002; 2002US-0406640P.
 PR 29-AUG-2002; 2002US-0406642P.
 PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.
 PR 29-AUG-2002; 2002US-0406655P.
 PR 29-AUG-2002; 2002US-0406666P.
 PR 17-SEP-2002; 2002US-0410946P.
 PR 17-SEP-2002; 2002US-0410947P.
 PR 17-SEP-2002; 2002US-0410948P.
 PR 17-SEP-2002; 2002US-0410953P.
 PR 17-SEP-2002; 2002US-0410957P.
 PR 17-SEP-2002; 2002US-0410958P.
 PR 17-SEP-2002; 2002US-0410959P.
 PR 17-SEP-2002; 2002US-0410960P.
 PR 17-SEP-2002; 2002US-0410961P.
 PR 17-SEP-2002; 2002US-0410962P.
 PR 17-SEP-2002; 2002US-0411019P.
 PR 17-SEP-2002; 2002US-0411022P.
 PR 17-SEP-2002; 2002US-0411023P.
 PR 17-SEP-2002; 2002US-0411024P.
 PR 17-SEP-2002; 2002US-0411032P.
 PR 17-SEP-2002; 2002US-0411035P.
 PR 17-SEP-2002; 2002US-0411037P.
 PR 17-SEP-2002; 2002US-0411041P.
 PR 17-SEP-2002; 2002US-0411045P.
 PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411048P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411073P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411101P.
 PR 17-SEP-2002; 2002US-0411111P.
 PR 18-APR-2003; 2003US-0463700P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 02-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 22-MAY-2003; 2003US-0471336P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.
 PR 09-JUN-2003; 2003US-0476609P.
 PR 08-JUL-2003; 2003US-0476641P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 08-JUL-2003; 2003US-0485224P.
 PR 08-JUL-2003; 2003US-0485325P.
 PR 14-JUL-2003; 2003US-0486446P.
 PR 15-JUL-2003; 2003US-0486480P.
 PR 15-JUL-2003; 2003US-0486891P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 XX
 XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
 PA
 XX
 PI Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;
 PI Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
 XX
 XX WPI; 2004-348438/32.
 DR
 XX
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.
 XX
 PS Claim 1; SEQ ID NO 1848; 428pp; English.
 XX
 CC The present invention relates to an isolated nucleic acid molecule

CC encoding a polypeptide which is believed to be cytostatic,
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPWEB and is not in the specification.
 XX
 SQ Sequence 149 AA;

Query Match 61.1%; Score 44; DB 8; Length 149;
 Best Local Similarity 62.5%; Pred. No. 1.5e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGYWLTIW 8
 | : ||| : |
 Db 95 CHEFWLTW 102

RESULT 14
 AAEE31487
 ID AAE31487 standard; protein; 166 AA.
 XX
 AC AAE31487;
 XX
 XX
 DT 24-FEB-2003 (first entry)
 XX
 DE Human butyrophilin 2/3 protein.
 XX
 KW Human; B7-H1.2 protein; Butyrophilin 2/3 protein; transplant rejection;
 KW immunological condition; graft-versus-host disease; allergy; asthma;
 KW inflammatory bowel disease; sepsis; Alzheimer's disease; atherosclerosis;
 KW T-cell mediated inflammation; autoimmune disease; multiple sclerosis;
 KW systemic lupus erythematosus; autoimmune demyelination; Grave's disease;
 KW psoriasis; autoimmune diabetes; diabetic neuropathy; HIV infection;
 KW rheumatoid arthritis; human immunodeficiency virus; immunosuppressive;
 KW gene therapy; infection; virucide.
 XX
 OS Homo sapiens.
 XX
 PN WO200279474-A2.
 XX
 XX 10-OCT-2002.
 XX
 XX 08-JAN-2002; 2002WO-US000590.
 XX
 PF 08-JAN-2001; 2001US-0260617P.
 PR 19-JAN-2001; 2001US-0262737P.
 PR 07-JAN-2002; 2002US-00260617.
 XX
 XX (IMMV) IMMUNEX CORP.
 XX
 XX Baum PR, Dubose RF, Wiley SR;
 XX WPI; 2003-046816/04.
 XX
 DR New B7-H1.2 or Butyrophilin 2/3 polypeptide of the human B7 polypeptide
 FT family, useful for treating an immunological condition e.g. transplant
 PT rejection.
 PT
 XX Example 1; Page 64-65; 99pp; English.
 PS
 XX The invention relates to B7-H1.2 or Butyrophilin 2/3 polypeptides of
 CC human B7 polypeptide family and polynucleotides encoding such proteins.
 CC Sequences of the invention are useful for treating an immunological
 CC conditions (e.g., transplant rejection, graft-versus-host disease,
 CC allergy, asthma, inflammatory bowel disease, sepsis), diseases that are
 CC caused or exacerbated by T-cell mediated inflammation (e.g., Alzheimer's
 CC disease or atherosclerosis), autoimmune diseases (e.g., systemic lupus
 CC erythematosus, autoimmune demyelination, Grave's disease, psoriasis,
 CC multiple sclerosis, autoimmune diabetes, diabetic neuropathy, rheumatoid
 CC arthritis), bacterial or viral infections such as human immunodeficiency
 CC virus (HIV) infection, delayed reconstitution of T-cells, defects in T-

CC cell or accessory cell function or congenital immunodeficiencies. They
 CC are also used in gene therapy. The present sequence is human butyrophilin
 CC consensus protein
 XX
 SQ Sequence 166 AA;

Query Match 61.1%; Score 44; DB 6; Length 166;
 Best Local Similarity 62.5%; Pred. No. 1.6e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 YWLTWGC 10
 | : ||| : |
 Db 127 YWEVWGC 134

RESULT 15
 AAU91293
 ID AAU91293 standard; protein; 415 AA.
 XX
 AC AAU91293;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Human NOV8 protein.
 XX
 KW Human; NOVX; gene therapy; cardiomyopathy; atherosclerosis; diabetes;
 KW cell signal processing; metabolic pathway modulation; inflammation;
 KW autoimmune disorder; scleroderma; transplantation; allergy;
 KW systemic lupus erythematosus; haemophilia; Alzheimer's disease;
 KW graft versus host disease; Leech-Nyhan syndrome; periodontitis;
 KW pancreatitis; musculoskeletal disorder; Parkinson's disease;
 KW Huntington's disease; behavioural disorder; pain; obesity; wound healing;
 KW neurodegenerative disorder; neuropsychiatric disorder; hypertension;
 KW growth disorder; reproductive disorder; lung disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200216600-A2.
 XX
 XX 28-FEB-2002.
 XX
 XX 27-AUG-2001; 2001WO-US026518.
 XX
 PR 25-AUG-2000; 2000US-0227800P.
 PR 25-AUG-2000; 2000US-0228205P.
 PR 25-AUG-2000; 2000US-0228324P.
 PR 30-AUG-2000; 2000US-0228997P.
 PR 30-AUG-2000; 2000US-0229185P.
 PR 01-SEP-2000; 2000US-0229780P.
 PR 01-SEP-2000; 2000US-0229848P.
 PR 01-SEP-2000; 2000US-0229850P.
 PR 22-JAN-2001; 2001US-0263337P.
 PR 31-JAN-2001; 2001US-0265518P.
 PR 15-MAR-2001; 2001US-0276451P.
 PR 27-MAR-2001; 2001US-0279196P.
 PR 24-AUG-2001; 2001US-00939398.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Gerlach V, Macdougall JR, Smithson G, Stone DJ, Ellerman K;
 PI Spytek KA, Zerhusen BD, Rastelli L, Verney CAM, Patturajan M;
 PI Tchernev VT, Padigar M, Taupier RJ;
 XX
 DR WPI; 2002-292064/33.
 DR N-PSDB; ABKS5577.
 XX
 XX New isolated cytoplasmic, nuclear, membrane bound and secreted
 PT polypeptides, termed NOVX, useful for treating inflammation, autoimmune
 PT disorders, hemophilia, Leech-Nyhan syndrome, pancreatitis,
 XX musculoskeletal disorders.
 PS Claim 1; Page 108; 245pp; English.
 XX

CC The invention relates to an isolated cytoplasmic, nuclear, membrane bound
 CC or secreted polypeptide, designated NOVX (actually NOV1, 2a, 2b, 3a, 3b,
 CC 4, 5a, 5b, 5c, 5d, 5e, 5f, 5g, 5h, 5i, 6, 7 and 8), a variant of NOVX, a
 CC mature form, or a variant of the mature form of NOVX. Also included are a
 CC polynucleotide encoding NOVX (or its complement), a vector comprising the
 CC polynucleotide, a cell comprising the vector, an anti-NOVX antibody,
 CC determining the presence of NOVX in a sample using the antibody,
 CC determining the presence of NOVX polynucleotide in a sample which binds to
 CC which binds to NOVX polynucleotide, identifying an agent which binds to
 CC NOVX (including modulators of NOVX). NOVX, the polynucleotide and the
 CC antibody are useful for diagnosing, treating or preventing a NOVX-
 CC associated disorder selected from cardiomyopathy, atherosclerosis,
 CC diabetes, a disorder related to cell signal processing and metabolic
 CC pathway modulation, inflammation, autoimmune disorders, scleroderma,
 CC transplantation, allergies, systemic lupus erythematosus, haemophilia,
 CC graft-versus host disease, Alzheimer's disease, stroke, Lesch-Nyhan
 CC syndrome, periodontitis, pancreatitis, musculoskeletal disorders,
 CC Parkinson's disease, Huntington's disease, behavioural disorders, pain,
 CC neurodegenerative and neuropsychiatric disorders, hypertension, wound
 CC healing, obesity, growth and reproductive disorders, lung diseases and
 CC many other diseases and disorders listed in the specification. NOVX, the
 CC polynucleotide and the antibody are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g., diagnostic assays, prognostic assays,
 CC monitoring clinical trials and pharmacogenomic), and in methods of
 CC treatment (e.g., therapeutic and prophylactic). NOVX is useful as
 CC immunogen to produce antibodies immunospecific for NOVX, as vaccines to
 CC screen for potential agonist and antagonist compounds, and as bait
 CC protein in a two-hybrid or three-hybrid assay. The polynucleotide is
 CC useful in gene therapy, to express NOVX, to detect NOVX mRNA or a genetic
 CC lesion in a NOVX gene, and to modulate NOVX activity. The vector is
 CC useful for producing non-human transgenic animals. The antibody is useful
 CC for isolating, and purifying NOVX and to monitor protein levels in tissue
 CC as part of a clinical testing procedure. The present sequence represents
 CC a NOVX protein
 CC
 XX

SQ Sequence 415 AA;

Query Match 61.1%; Score 44; DB 5; Length 415;

Best Local Similarity 62.5%; Pred. No. 3.5e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGYWLTIW 8

| : ||| : |

Db 95 CHEFLLTVW 102

Search completed: December 30, 2004, 13:07:53

Job time : 72.6226 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 12:59:07 ; Search time 23.0189 Seconds
(without alignments)
28.810 Million cell updates/sec

Title: US-10-046-922-35

Perfect score: 72

Sequence: 1 CGYWLTIWGC 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44.5	61.8	1024	4	US-09-562-737-84 Sequence 84, Appl
2	43	59.7	492	4	US-09-107-532A-6945 Sequence 6945, Ap
3	42.5	59.0	1024	4	US-09-562-737-86 Sequence 86, Appl
4	42.5	59.0	1765	4	US-09-354-147C-2 Sequence 2, Appli
5	42.5	59.0	1765	4	US-09-354-147C-3 Sequence 3, Appli
6	42	58.3	161	4	US-09-325-932A-187 Sequence 187, App
7	42	58.3	273	4	US-09-270-767-46926 Sequence 46926, A
8	42	58.3	1498	4	US-09-792-616-9 Sequence 9, Appli
9	42	58.3	1503	4	US-09-792-616-3 Sequence 3, Appli
10	41	56.9	222	4	US-09-071-035-408 Sequence 408, App
11	41	56.9	229	4	US-09-134-000C-3630 Sequence 3630, Ap
12	41	56.9	264	4	US-09-540-236-2978 Sequence 2978, Ap
13	41	56.9	266	4	US-09-071-035-406 Sequence 406, App
14	40.5	56.2	1024	4	US-09-562-737-87 Sequence 87, Appl
15	40.5	56.2	1024	4	US-09-562-737-81 Sequence 81, Appl
16	40.5	56.2	1835	3	US-08-836-325-15 Sequence 15, Appl
17	40.5	56.2	1835	3	US-09-457-571-15 Sequence 15, Appl
18	40.5	56.2	1836	4	US-10-162-012-24 Sequence 24, Appl
19	40.5	56.2	1969	3	US-08-836-325-16 Sequence 16, Appl
20	40.5	56.2	1969	4	US-09-457-571-16 Sequence 16, Appl
21	40.5	56.2	1976	3	US-09-024-020B-9 Sequence 9, Appli
22	40.5	56.2	1976	3	US-09-425-043-9 Sequence 9, Appli
23	40.5	56.2	1977	4	US-09-976-594-757 Sequence 757, App
24	40.5	56.2	1977	4	US-09-919-039-367 Sequence 367, App
25	40.5	56.2	1978	3	US-09-024-020B-3 Sequence 3, Appli
26	40.5	56.2	1978	3	US-09-425-043-3 Sequence 3, Appli
27	40.5	56.2	1984	3	US-08-836-325-10 Sequence 10, Appl

28	40.5	56.2	1984	4	US-09-457-571-10 Sequence 10, Appl
29	40.5	56.2	1988	3	US-09-024-020B-4 Sequence 4, Appli
30	40.5	56.2	1988	3	US-09-425-043-4 Sequence 4, Appli
31	40.5	56.2	1989	3	US-08-836-325-11 Sequence 11, Appl
32	40.5	56.2	1989	3	US-08-836-325-12 Sequence 12, Appl
33	40.5	56.2	1989	4	US-09-457-571-11 Sequence 11, Appl
34	40.5	56.2	1989	4	US-09-457-571-12 Sequence 12, Appl
35	40.5	56.2	2005	3	US-08-836-325-7 Sequence 7, Appli
36	40.5	56.2	2005	4	US-09-457-571-7 Sequence 7, Appli
37	40.5	56.2	2016	3	US-09-634-920-4 Sequence 4, Appli
38	40.5	56.2	2016	4	US-09-514-907A-2 Sequence 2, Appli
39	40.5	56.2	2016	4	US-09-896-994-2 Sequence 2, Appli
40	40.5	56.2	2016	4	US-09-840-125-4 Sequence 4, Appli
41	40	55.6	362	1	US-08-415-571-6 Sequence 6, Appli
42	40	55.6	668	4	US-09-248-796A-19350 Sequence 19350, A
43	39.5	54.9	73	4	US-09-513-999C-4541 Sequence 4541, Ap
44	39	54.2	430	4	US-09-443-041A-18 Sequence 18, Appl
45	39	54.2	450	4	US-09-443-041A-22 Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-09-562-737-84
; Sequence 84, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-84

Query Match 61.8%; Score 44.5; DB 4; Length 1024;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;
QY 1 CGYWL-TIWGC 10
|||:|:|:
DB 892 CGHWIETWDC 902

RESULT 2
US-09-107-532A-6945
; Sequence 6945, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FASCIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC

```
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107.532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6945:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 492 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...492
; SEQUENCE DESCRIPTION: SEQ ID NO: 6945:
US-09-107-532A-6945
Query Match 59.7%; Score 43; DB 4; Length 492;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GYWLTIW 8
Db 112 GYWLTCW 118

RESULT 3
US-09-562-737-86
; Sequence 86, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562.737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-86
Query Match 59.0%; Score 42.5; DB 4; Length 1024;
Best Local Similarity 63.6%; Pred. No. 3.1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Qy 1 CGYWL-TIWGC 10
Db 892 CGEWLETMWDC 902

RESULT 4
US-09-354-147C-2
; Sequence 2, Application US/09354147C
; Patent No. 6573067
; GENERAL INFORMATION:
; APPLICANT: Dib-Hajj, Sulayman
; TITLE OF INVENTION: Modulation of Sodium Channels in Dorsal Root Ganglia
; FILE REFERENCE: 44574-5004-01-US
; CURRENT APPLICATION NUMBER: US/09/354.147C
; CURRENT FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US 60/072,990
; PRIOR FILING DATE: 1998-01-29
; PRIOR APPLICATION NUMBER: US 60/109,402
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: PCT/US99/02008
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1765
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (652)..(1334)
; OTHER INFORMATION: Xaa at position 652 is Leu; Xaa at position 1334 is Asn
; OTHER INFORMATION: or Lys. Xaa's result from n's in SEQ ID NO: 1.
US-09-354-147C-2
Query Match 59.0%; Score 42.5; DB 4; Length 1765;
Best Local Similarity 54.5%; Pred. No. 5.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 CGYWL-TIWGC 10
Db 754 CGEWIENMWC 764

RESULT 5
US-09-354-147C-3
; Sequence 3, Application US/09354147C
; Patent No. 6573067
; GENERAL INFORMATION:
; APPLICANT: Dib-Hajj, Sulayman
; TITLE OF INVENTION: Modulation of Sodium Channels in Dorsal Root Ganglia
; FILE REFERENCE: 44574-5004-01-US
; CURRENT APPLICATION NUMBER: US/09/354.147C
; CURRENT FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US 60/072,990
; PRIOR FILING DATE: 1998-01-29
; PRIOR APPLICATION NUMBER: US 60/109,402
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: PCT/US99/02008
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1765
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: putative amino acid seq. of rat Nan
US-09-354-147C-3
Query Match 59.0%; Score 42.5; DB 4; Length 1765;
Best Local Similarity 54.5%; Pred. No. 5.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 CGYWL-TIWGC 10
Db 754 CGEWIENMWC 764
```



```
RESULT 6
US-09-325-932A-187
; Sequence 187, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE OF INVENTION: death and their use in the modification of forestry plant develop
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 187
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-325-932A-187

Query Match      58.3%; Score 42; DB 4; Length 161;
Best Local Similarity 66.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      2 GYWLTIWGC 10
      |||| : ||
Db      100 GYWLNLSC 108

RESULT 7
US-09-270-767-46926
; Sequence 46926, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46926
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46926

Query Match      58.3%; Score 42; DB 4; Length 273;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CGYWLTIW 8
      |||| : ||
Db      266 CGYWTATF 273

RESULT 8
US-09-792-616-9
; Sequence 9, Application US/09792616
; Patent No. 6780587
; GENERAL INFORMATION:
; APPLICANT: PXE International, Inc.
; APPLICANT: University of Hawaii
; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (MRP6) causing
; FILE OF INVENTION: Pseudoxanthoma Elasticum
; FILE REFERENCE: PXE-001
; CURRENT APPLICATION NUMBER: US/09/792,616
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1498
```

```
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-792-616-9

Query Match      58.3%; Score 42; DB 4; Length 1498;
Best Local Similarity 71.4%; Pred. No. 5.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GYWLTIW 8
      |||| : ||
Db      960 GYWLSLW 966

RESULT 9
US-09-792-616-3
; Sequence 3, Application US/09792616
; Patent No. 6780587
; GENERAL INFORMATION:
; APPLICANT: PXE International, Inc.
; APPLICANT: University of Hawaii
; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (MRP6) causing
; FILE OF INVENTION: Pseudoxanthoma Elasticum
; FILE REFERENCE: PXE-001
; CURRENT APPLICATION NUMBER: US/09/792,616
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1503
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-616-3

Query Match      58.3%; Score 42; DB 4; Length 1503;
Best Local Similarity 71.4%; Pred. No. 5.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GYWLTIW 8
      |||| : ||
Db      965 GYWLSLW 971

RESULT 10
US-09-071-035-408
; Sequence 408, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 408:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-408

Query Match 56.9%; Score 41; DB 4; Length 222;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYWLTIW 9
DB 179 GTWITLWG 186

RESULT 11
US-09-134-000C-3630
; Sequence 3630, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3630
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3630

Query Match 56.9%; Score 41; DB 4; Length 229;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYWLTIW 9
DB 203 GTWITLWG 210

RESULT 12
US-09-540-236-2978
; Sequence 2978, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2978
; LENGTH: 264
; TYPE: PRT
; ORGANISM: M catarrhalis
US-09-540-236-2978

Query Match 56.9%; Score 41; DB 4; Length 264;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 408:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-406

Query Match 56.9%; Score 41; DB 4; Length 266;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYWLTIW 9
DB 203 GTWITLWG 210

RESULT 14
US-09-562-737-81
; Sequence 81, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 1024
; TYPE: PRT
```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-81

Query Match 56.2%; Score 40.5; DB 4; Length 1024;
Best Local Similarity 54.5%; Pred. No. 5.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10
|||:|:|
Db 892 CGEWIGTWDC 902

RESULT 15
US-09-562-737-87
; Sequence 87, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-87

Query Match 56.2%; Score 40.5; DB 4; Length 1024;
Best Local Similarity 54.5%; Pred. No. 5.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10
|||:|:|
Db 892 CGEWIGTWDC 902

Search completed: December 30, 2004, 13:19:20
Job time : 24.0189 secs

THIS PAGE BLANK (USP 10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 13:08:04 ; Search time 74.5283 Seconds
(without alignments)
48.267 Million cell updates/sec

Title: US-10-046-922-35
Perfect score: 72
Sequence: 1 CGYWLTIWGC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues
Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	100.0	10	13	US-10-046-922-35
2	54	75.0	10	13	Sequence 35, Appl
3	50	69.4	10	13	Sequence 34, Appl
4	47.5	66.0	304	14	US-10-046-922-73
5	47.5	66.0	304	14	Sequence 2288, Ap
6	47	65.3	25	14	Sequence 1934, Ap
7	47	65.3	63	16	US-10-280-066-476
8	46	63.9	17	14	Sequence 125253,
9	46	63.9	17	15	US-10-125-869A-54
10	46	63.9	136	15	Sequence 54, Appl
11	45	62.5	828	15	Sequence 278, App
12	44.5	61.8	1024	14	Sequence 743, App
13	44	61.1	103	16	US-10-112-944-743
					Sequence 181150,

14	61.1	161	16	US-10-767-701-54778	Sequence 54778, A	
15	59.7	13	14	US-10-125-869A-73	Sequence 73, Appl	
16	43	59.7	13	US-10-462-262-297	Sequence 297, Appl	
17	43	59.7	291	14	US-10-369-493-19313	Sequence 19313, A
18	43	59.7	454	14	US-10-032-585-7816	Sequence 7816, Ap
19	43	59.7	459	16	US-10-437-963-160191	Sequence 160191, A
20	43	59.7	474	15	US-10-382-122A-57928	Sequence 57928, A
21	43	59.7	671	16	US-10-437-963-190740	Sequence 190740, A
22	42.5	59.0	152	11	US-09-833-245-1533	Sequence 1533, Ap
23	42.5	59.0	152	15	US-10-264-237-2457	Sequence 2457, Ap
24	42.5	59.0	159	11	US-09-833-245-1534	Sequence 1534, Ap
25	42.5	59.0	161	15	US-10-276-774-2376	Sequence 2376, Ap
26	42.5	59.0	161	15	US-10-296-115-1337	Sequence 1337, Ap
27	42.5	59.0	176	14	US-10-161-927-10	Sequence 10, Appl
28	42.5	59.0	1024	14	US-10-211-962-86	Sequence 86, Appl
29	42.5	59.0	1765	14	US-10-388-470-2	Sequence 2, Appl
30	42.5	59.0	1765	14	US-10-388-470-3	Sequence 3, Appl
31	42	58.3	90	11	US-09-864-408A-8146	Sequence 8146, Ap
32	42	58.3	161	14	US-10-219-220-187	Sequence 187, App
33	42	58.3	1498	10	US-09-792-616-9	Sequence 9, Appl
34	42	58.3	1498	16	US-10-764-328-9	Sequence 9, Appl
35	42	58.3	1503	10	US-09-792-616-3	Sequence 3, Appl
36	42	58.3	1503	16	US-10-764-328-3	Sequence 3, Appl
37	41.5	57.6	122	15	US-10-108-260A-2641	Sequence 2641, Ap
38	41.5	57.6	620	14	US-10-369-493-119	Sequence 119, App
39	41	56.9	14	14	US-10-125-869A-114	Sequence 114, App
40	41	56.9	14	15	US-10-462-262-338	Sequence 338, App
41	41	56.9	34	9	US-09-864-761-43458	Sequence 43458, A
42	41	56.9	57	15	US-10-424-599-179308	Sequence 179308,
43	41	56.9	78	17	US-10-425-115-204642	Sequence 204642,
44	41	56.9	82	15	US-10-424-599-240891	Sequence 240891,
45	41	56.9	104	17	US-10-425-115-332065	Sequence 332065,

ALIGNMENTS

RESULT 1
US-10-046-922-35
; Sequence 35, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046.922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35
; LENGTH: 10
; TYPE: PRT
; ORGANISM: isolated peptide
US-10-046-922-35

Query Match 100.0%; Score 72; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGYWLTIWGC 10
|||||
DB 1 CGYWLTIWGC 10

RESULT 2
US-10-046-922-34
; Sequence 34, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki

; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046.922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 10
; TYPE: PRT
; ORGANISM: isolated peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(1)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (10)..(10)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-34

Query Match 75.0%; Score 54; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GYWLTIWG 9
| | | | | | | |
Db 2 GYWLTIWG 9

RESULT 3
US-10-046-922-73
; Sequence 73, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046.922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73
; LENGTH: 10
; TYPE: PRT
; ORGANISM: peptide library
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)..(7)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (9)..(9)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-73

Query Match 69.4%; Score 50; DB 13; Length 10;
Best Local Similarity 60.0%; Pred. No. 2.5;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGYWLTIWGC 10
| | | | | | | |
Db 1 CGYWLTIWGC 10

RESULT 4
US-10-017-161-2288
; Sequence 2288, Application US/10017161
; Publication No. US20030143669A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA

; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017.161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2288
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (73)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (79)..(83)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (85)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (89)..(91)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (96)..(97)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (101)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (107)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (118)..(119)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (121)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (123)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (162)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (178)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (194)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (210)..(211)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (219)
; OTHER INFORMATION: Variable amino acid

```
;
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (256)..(289)
; OTHER INFORMATION: Variable amino acid
US-10-017-161-2288

Query Match      66.0%; Score 47.5; DB 14; Length 304;
Best Local Similarity 60.0%; Pred. No. 88;
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy      1 CGYWLTIWGC 10
      |||:|:|
Db      108 CGFW-AVWGC 116

RESULT 5
US-10-292-798-1934
; Sequence 1934, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1934
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (73)..(73)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (80)..(83)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (89)..(91)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (96)..(97)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (101)..(101)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (107)..(107)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (116)..(119)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (121)..(121)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (123)..(123)
```

```
;
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (178)..(178)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (210)..(211)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (215)..(219)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (256)..(289)
; OTHER INFORMATION: Variable amino acid
US-10-292-798-1934

Query Match      66.0%; Score 47.5; DB 14; Length 304;
Best Local Similarity 60.0%; Pred. No. 88;
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy      1 CGYWLTIWGC 10
      |||:|:|
Db      108 CGFW-AVWGC 116

RESULT 6
US-10-280-066-476
; Sequence 476, Application US/10280066
; Publication No. US20030180718A1
; GENERAL INFORMATION:
; APPLICANT: Pillutla, Renuka C.
; APPLICANT: Briessette, Renee
; APPLICANT: Spruyt, Michael
; APPLICANT: Dedova, Olga
; APPLICANT: Blume, Arthur J.
; APPLICANT: Prendergast, John
; APPLICANT: Goldstein, Neil I.
; TITLE OF INVENTION: TARGET SPECIFIC SCREENING AND ITS USE FOR IDENTIFYING TARGET BIND
; FILE REFERENCE: 2598-4009US1
; CURRENT APPLICATION NUMBER: US/10/280,066
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 60/345,471
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 537
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 476
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Eschericia coli
; FEATURE:
; NAME/KEY: MISC_FEATURE
; OTHER INFORMATION: T1el-20C-3-D116
US-10-280-066-476

Query Match      65.3%; Score 47; DB 14; Length 25;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 CGYWLTIW 9
      ||||:|
Db      5 CGYWGELWG 13

RESULT 7
US-10-437-963-125253
; Sequence 125253, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
```

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 125253
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_27915C.1.pep
US-10-437-963-125253

Query Match      65.3%; Score 47; DB 16; Length 63;
Best Local Similarity 60.0%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGYWLTIWGC 10
Db 37 CGHYLKAWGC 46

RESULT 8
US-10-125-869A-54
; Sequence 54, Application US/10125869A
; Publication No. US20030199671A1
; GENERAL INFORMATION:
; APPLICANT: Rondon, Isaac Jesus
; APPLICANT: Wu, Qi-Long
; APPLICANT: Ley, Arthur C.
; APPLICANT: Stochl, Mark
; APPLICANT: Ransohoff, Thomas C.
; APPLICANT: Potter, M. Daniel (deceased)
; TITLE OF INVENTION: BINDING MOLECULES FOR FC-REGION
; FILE REFERENCE: 3421.1006-001
; CURRENT APPLICATION NUMBER: US/10/125,869A
; CURRENT FILING DATE: 2002-11-19
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fc region binding polypeptide
US-10-125-869A-54

Query Match      63.9%; Score 46; DB 14; Length 17;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGYWLTIWGC 9
Db 4 CGFWPRWIG 12

RESULT 9
US-10-462-262-278
; Sequence 278, Application US/10462262
; Publication No. US20040009534A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Aaron K.
; APPLICANT: Dawson, Bruce M.
```

```
; TITLE OF INVENTION: PROTEIN ANALYSIS
; FILE REFERENCE: 10280-052001
; CURRENT APPLICATION NUMBER: US/10/462,262
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/388,642
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 278
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: immunoglobulin binding polypeptide
US-10-462-262-278

Query Match      63.9%; Score 46; DB 15; Length 17;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGYWLTIWGC 9
Db 4 CGFWPRWIG 12

RESULT 10
US-10-112-944-743
; Sequence 743, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Weng, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 743
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(136)
; OTHER INFORMATION: Xaa = any amino acid or symbol as shown in the table 8 as set for
```



```
; OTHER INFORMATION: in Example 2
US-10-112-944-743
Query Match      63.9%; Score 46; DB 15; Length 136;
Best Local Similarity 60.0%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGYWLTIWGC 10
   |||:||||
Db 93 CGRWDLWGC 102

RESULT 11
US-10-282-122A-77862
; Sequence 77862, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77862
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Yersinia pestis
US-10-282-122A-77862

Query Match      62.5%; Score 45; DB 15; Length 828;
Best Local Similarity 58.3%; Pred. No. 43e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 CGYWLTI--WGC 10
   |||:||||
Db 372 CGLWLELLSWGC 383

RESULT 12
US-10-211-962-84
```

```
; Sequence 84, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562,737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 84
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-211-962-84

Query Match      61.8%; Score 44.5; DB 14; Length 1024;
Best Local Similarity 54.5%; Pred. No. 6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10
   |||:||||
Db 892 CGHWIETWDC 902

RESULT 13
US-10-437-963-181150
; Sequence 181150, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 181150
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_78451C.1.pep
US-10-437-963-181150

Query Match      61.1%; Score 44; DB 16; Length 103;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 YWLTIWGC 10
   :|:|||||
Db 19 WWSVWGC 26

RESULT 14
US-10-767-701-54778
; Sequence 54778, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
```

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 54778
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 14593586.pap
US-10-767-701-54778

Query Match 61.1%; Score 44; DB 16; Length 161;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGYWLTIW 8
||| |
Db 150 CGYWLKEW 157

RESULT 15
US-10-125-869A-73
; Sequence 73, Application US/10125869A
; Publication No. US20030199671A1
; GENERAL INFORMATION:
; APPLICANT: Rondon, Isaac Jesus
; APPLICANT: Wu, Qi-Long
; APPLICANT: Ley, Arthur C.
; APPLICANT: Stochl, Mark
; APPLICANT: Ransohoff, Thomas C.
; APPLICANT: Potter, M. Daniel (deceased)
; TITLE OF INVENTION: BINDING MOLECULES FOR FC-REGION
; FILE REFERENCE: 3421.1006-001
; CURRENT APPLICATION NUMBER: US/10/125,869A
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/284,534
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fc region binding polypeptide
US-10-125-869A-73

Query Match 59.7%; Score 43; DB 14; Length 13;
Best Local Similarity 62.5%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GYWLTIWG 9
||| :||
Db 1 GYWCNVWG 8

Search completed: December 30, 2004, 13:50:01
Job time : 75.5283 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 12:58:17 ; Search time 11.5094 Seconds
(without alignments)
83.598 Million cell updates/sec

Title: US-10-046-922-35

Perfect score: 72

Sequence: 1 CGYWLTTWGC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	63.9	218	2 S76385	hypothetical prote
2	45	62.5	828	2 AD0412	ATP-dependent heli
3	43.5	60.4	1993	2 T30902	sodium channel sca
4	42.5	59.0	1765	2 T42388	sodium channel alp
5	42	58.3	72	2 S30980	gene 35 protein -
6	42	58.3	441	2 C95307	probable transport
7	42	58.3	1502	2 T42216	multidrug resistan
8	41.5	57.6	620	2 F70439	oxalosuccinate decar
9	41	56.9	307	2 T48166	hypothetical prote
10	41	56.9	391	2 P41117	replication protei
11	41	56.9	419	2 E90446	permease [imported
12	41	56.9	739	2 T29407	hypothetical prote
13	40.5	56.2	200	2 I48108	sodium channel alp
14	40.5	56.2	1681	2 A55138	sodium channel mNa
15	40.5	56.2	1682	2 A45380	sodium channel pro
16	40.5	56.2	1820	1 CHEE	sodium channel pro
17	40.5	56.2	1835	2 I54323	sodium channel alp
18	40.5	56.2	1836	2 I64893	sodium channel alp
19	40.5	56.2	1836	2 J50648	sodium channel alp
20	40.5	56.2	1836	2 I51964	sodium channel alp
21	40.5	56.2	1840	1 CHRTM1	sodium channel pro
22	40.5	56.2	1951	2 S00320	sodium channel pro
23	40.5	56.2	1976	2 I56555	sodium channel pro
24	40.5	56.2	1977	2 S54771	sodium channel alp
25	40.5	56.2	1983	2 A60054	sodium channel pro
26	40.5	56.2	2005	2 A46269	sodium channel alp
27	40.5	56.2	2005	2 B25019	sodium channel pro
28	40.5	56.2	2009	2 A25019	sodium channel pro
29	40.5	56.2	2016	2 A38195	sodium channel pro

30 40.5 56.2 2019 2 A33996 sodium channel pro
31 40.5 56.2 2049 2 T43161 sodium channel pro
32 40 55.6 142 2 C34903 ig heavy chain pre
33 40 55.6 339 2 F97190 phenylalanyl-CRNA
34 40 55.6 359 2 F95406 probable ABC trans
35 40 55.6 425 2 B71038 probable Na+/H+-ex
36 40 55.6 508 2 C35282 probable ABC trans
37 40 55.6 2344 2 S64740 genome polyprotein
38 39.5 54.9 298 2 AH0289 probable aldo/keto
39 39.5 54.9 345 1 JH0185 D-amino-acid oxida
40 39.5 54.9 347 1 OXP6DA D-amino-acid oxida
41 39.5 54.9 347 1 S01340 D-amino-acid oxida
42 39.5 54.9 347 1 JX0132 D-amino-acid oxida
43 39 54.2 376 2 AF1978 hypothetical prote
44 39 54.2 392 2 A53580 neurexin III beta
45 39 54.2 426 2 B53580 neurexin III beta

ALIGNMENTS

RESULT 1

S76385

hypothetical protein - Synechocystis sp. (strain PCC 6803)

C/Species: Synechocystis sp.

A/Variety: PCC 6803

C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C/Accession: S76385

R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.,

O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

S.

A/Reference number: S74322; MUID:97061201; PMID:8905231

A/Accession: S76385

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-218 <KAN>

A/Cross-references: UNIPROT:Q55705; EMBL:D64000; GB:AB001339; NID:G1001484; PIDN:BAA1023

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C/Superfamily: Probable alkaline phosphatase yncG

Query Match 63.9%; Score 46; DB 2; Length 218;
Best Local Similarity 66.7%; Pred. No. 8.1;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGYWLTTWG 9

Db 73 CGYWVGRWG 81

RESULT 2

AD0412

ATP-dependent helicase [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis

C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C/Accession: AD0412

R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: AD0412

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-828 <KUR>

A/Cross-references: UNIPROT:Q82BL4; GB:AL590842; PIDN:CAC92624.1; PID:G15981320; GSPDB:C

C/Genetics:

A/Gene: hrpB

C/Superfamily: ATP-dependent RNA helicase, HrpB type

Query Match 62.5%; Score 45; DB 2; Length 828;

Best Local Similarity 58.3%; Pred. No. 36;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 CGYWLTI--WGC 10
||| : |||
Db 372 CGLWLELLSWGC 383

RESULT 3

T30902
sodium channel SCAP1 alpha chain - California sea hare
C;Species: Aplysia californica (California sea hare)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30902
R;Dyer, J.R.; Johnston, W.L.; Castellucci, V.F.; Dunn, R.J.
DNA Cell Biol. 16, 347-356, 1997
A;Title: Cloning and tissue distribution of the Aplysia Na+ channel alpha-subunit cDNA.
A;Reference number: Z20929; MUID:97238630; PMID:9115644
A;Accession: T30902
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1993 <DY>
A;Cross-references: UNIPROT:P90670; EMBL:U66915; NID:g1842248; PID:g1842249; PIDN:AAC474
C;Superfamily: sodium channel protein

Query Match 60.4%; Score 43.5; DB 2; Length 1993;
Best Local Similarity 54.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10
||| : |||
Db 942 CGEWIESWGC 952

RESULT 4

T42388
sodium channel alpha chain - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T42388
R;Dib-Hajj, S.D.; Tyrell, L.; Black, J.A.; Waxman, S.G.
Proc. Natl. Acad. Sci. U.S.A. 95, 8963-8968, 1998
A;Title: NaN, a novel voltage-gated Na channel, is expressed preferentially in peripheral
A;Reference number: Z22149; MUID:98338024; PMID:9671787
A;Accession: T42388
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1765 <DIB>
A;Cross-references: UNIPROT:O88457; EMBL:AF059030; NID:g3372614; PID:g3372615; PIDN:AAC4
A;Experimental source: strain Sprague-Dawley; dorsal root ganglia
A;Note: preferentially expressed in sensory neurons within dorsal root ganglia and trigem
C;Superfamily: sodium channel protein

Query Match 59.0%; Score 42.5; DB 2; Length 1765;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10
||| : |||
Db 754 CGEWIENWGC 764

RESULT 5

S30980
gene 35 protein - Mycobacterium phage L5
C;Species: Mycobacterium phage L5
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: S30980
R;Donnelly-Wu, M.K.; Jacobs Jr., W.R.; Hatfull, G.F.
Mol. Microbiol. 7, 407-417, 1993
A;Title: Superinfection immunity of mycobacteriophage L5: applications for genetic trans
A;Reference number: S30949; MUID:93211283; PMID:8459767
A;Accession: S30980

A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-72 <DON>
A;Cross-references: UNIPROT:Q05245; EMBL:Z18946; NID:g15859; PIDN:CAA79411.1; PID:g15891
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1992
C;Genetics:
A;Gene: 35

Query Match 58.3%; Score 42; DB 2; Length 72;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYWLTIW 8
||| : |||
Db 50 CGMWLPWV 57

RESULT 6

C95307
probable transport protein SMA0684 [imported] - Sinorhizobium meliloti (strain 1021) mag
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: C95307
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowse
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.;
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: C95307
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-441 <KUR>
A;Cross-references: UNIPROT:Q92ZT6; GB:AE006469; PIDN:AAK5021.1; PID:g14523451; GSPDB:G
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.;
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMA0684
A;Genome: plasmid
C;Superfamily: L-lysine transport protein

Query Match 58.3%; Score 42; DB 2; Length 441;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GYWLTIW 8
||| : |||
Db 92 GYWISW 98

RESULT 7

T42216
multidrug resistance-associated protein homolog MLP-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T42216
R;Hirohashi, T.; Suzuki, H.; Ito, K.; Ogawa, K.; Kume, K.; Shimizu, T.; Sugiyama, Y.
Mol. Pharmacol. 53, 1068-1075, 1998
A;Title: Hepatic expression of multidrug resistance-associated protein-like proteins mai
A;Reference number: Z22081; MUID:98279126; PMID:9614210
A;Accession: T42216
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1502 <HIR>
A;Cross-references: UNIPROT:O88269; EMBL:AB010466; NID:g3242457; PIDN:BAA28954.1; PID:g3
A;Experimental source: strain Sprague-Dawley; liver

C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 58.3%; Score 42; DB 2; Length 1502;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GYWLTIW 8
DB 964 GYWLTLW 970
|||||

RESULT 8
F70439
oxaloacetate decarboxylase alpha chain - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
A:Accession: F70439
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: F70439
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-620 <AQF>
A:Cross-references: UNIPROT:Q67544; GB:AE000747; NID:G2983944; PIDN:AAC07497.1; PID:G2983944
A:Experimental source: strain VF5
C:Genetics:
A:Gene: oadA
C:Superfamily: Klebsiella pneumoniae oxaloacetate decarboxylase alpha chain; lipoyl/biotin
F:540-613/Domain: lipoyl/biotin-binding homology <LPB>
F:579/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 57.6%; Score 41.5; DB 2; Length 620;
Best Local Similarity 60.0%; Pred. No. 93;
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 CQYW-LTIWG 9
DB 38 CGFWLEWVG 47
|||||

RESULT 9
T48166
hypothetical protein T1008.150 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
A:Accession: T48166
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Lemcke, K.; Mayer, K.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24486
A:Accession: T48166
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-307 <BEV>
A:Cross-references: UNIPROT:Q9M030; EMBL:AL1161746
A:Experimental source: cultivar Columbia; BAC clone T1008
C:Genetics:
A:Map position: 5
A:Introns: 31/3; 66/2; 114/3; 149/2; 232/2; 284/1
A:Note: T1008.150

Query Match 56.9%; Score 41; DB 2; Length 307;
Best Local Similarity 50.0%; Pred. No. 60;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CQYWLTIWGC 10
DB 186 CSFVVSINGC 195
|||||

RESULT 10

PC4117
replication protein homolog - Pyrococcus sp. (fragment)
N:Alternate names: hypothetical 391 protein
C:Species: Pyrococcus sp.
C:Date: 07-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 17-May-1996
A:Accession: PC4117
R:Raehid, N.; Morikawa, M.; Imanaka, T.
Gene 166, 139-143, 1995
A:Title: An abnormally acidic TATA-binding protein from a hyperthermophilic archaeon.
A:Reference number: JC4514; MUID:96105215; PMID:8529878
A:Accession: PC4117
A:Molecule type: DNA
A:Residues: 1-391 <RAS>
A:Cross-references: DDBJ:D50018

Query Match 56.9%; Score 41; DB 2; Length 391;
Best Local Similarity 85.7%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 YWLTIWG 9
DB 155 YWLTIWG 161
|||||

RESULT 11
E30446
permease [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
A:Accession: E30446
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: E30446
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-419 <KUR>
A:Cross-references: UNIPROT:Q97VB7; GB:AE006641; NID:G13816037; PIDN:AAK42828.1; GSPDB:G
C:Genetics:
A:Gene: SSO2718

Query Match 56.9%; Score 41; DB 2; Length 419;
Best Local Similarity 62.5%; Pred. No. 78;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYWLTIWG 9
DB 365 GFWETLWG 372
|||||

RESULT 12
T29407
hypothetical protein C16C8.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
A:Accession: T29407
R:Waterston, R.; Le, T.T.; Gattung, S.
submitted to the EMBL Data Library, November 1996
A:Description: The sequence of C. elegans cosmid C16C8.
A:Reference number: Z20617
A:Accession: T29407
A:Status: Preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-739 <WAT>
A:Cross-references: EMBL:U80452; PIDN:AA837863.1; GSPDB:GN00020; CESP:C16C8.2
A:Experimental source: strain Bristol N2; clone C16C8
C:Genetics:
A:Gene: CESP:C16C8.2
A:Map position: 2
A:Introns: 68/2; 179/3; 253/1; 275/3; 327/2; 365/3; 397/1; 428/2; 463/3; 629/1; 668/2; 7

C;Superfamily: myeloperoxidase; myeloperoxidase homology

Query Match 56.9%; Score 41; DB 2; Length 739;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGYWLTIWGC 10

DB 118 CGFWATIREC 127

RESULT 13

I48108

sodium channel alpha subunit - long-tailed hamster (fragment)

C;Species: Cricetus longicaudatus (long-tailed hamster)

C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004

C;Accession: I48108

R;Lalik, P.H.; Krafte, D.S.; Ciccarelli, R.B.

Am. J. Physiol. 264, 803-809, 1993

A;Title: Characterization of endogenous Sodium channel gene expressed in chinese hamster

A;Reference number: I48107

A;Accession: I48108

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-200 <RES>

A;Cross-references: UNIPROT:Q60464; GB:M87541; NID:g191069; PIDN:AAA36979.1; PID:g553840

C;Genetics:

A;Gene: chol

C;Superfamily: sodium channel protein

C;Keywords: duplication

Query Match 56.2%; Score 40.5; DB 2; Length 200;

Best Local Similarity 54.5%; Pred. No. 49;

Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10

DB 26 CGEWETIWDG 36

RESULT 14

A55138

sodium channel mna2.3, voltage-gated - mouse

C;Species: Mus musculus (house mouse)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C;Accession: A55138

R;Felipe, A.; Knittle, T.J.; Doyle, K.L.; Tamkun, M.M.

J. Biol. Chem. 269, 30125-30131, 1994

A;Title: Primary structure and differential expression during development and pregnancy

A;Reference number: A55138; MUID:95074002; PMID:7982916

A;Accession: A55138

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1681 <FEL>

A;Cross-references: UNIPROT:Q62467; GB:L36179; NID:g609544; PIDN:AAA66192.1; PID:g806397

C;Superfamily: sodium channel protein

Query Match

Best Local Similarity 56.2%; Score 40.5; DB 2; Length 1681;

Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10

DB 686 CGEWETIWDG 696

RESULT 15

A45380

sodium channel protein alpha chain hNav2.1 - human

C;Species: Homo sapiens (man)

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C;Accession: A45380

R;George Jr., A.L.; Knittle, T.J.; Tamkun, M.M.

Proc. Natl. Acad. Sci. U.S.A. 89, 4893-4897, 1992

A;Title: Molecular cloning of an atypical voltage-gated sodium channel expressed in human

A;Reference number: A45380; MUID:92279233; PMID:1317577

A;Accession: A45380

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-1682 <GEO>

A;Cross-references: UNIPROT:Q01118; GB:M91556; NID:g189046; PIDN:AAA59899.1; PID:g189047

A;Experimental source: heart

A;Note: sequence extracted from NCBI backbone (NCBIP:104344)

C;Superfamily: sodium channel protein

C;Keywords: glycoprotein; membrane protein; phosphoprotein; sodium channel; voltage-gate

Query Match

Best Local Similarity 56.2%; Score 40.5; DB 2; Length 1682;

Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGYWL-TIWGC 10

DB 686 CGEWETIWDG 696

Search completed: December 30, 2004, 13:17:13

Job time : 12.6761 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 12:57:52 ; Search time 92.2641 Seconds
(without alignments)
62.362 Million cell updates/sec

Title: US-10-046-922-35

Perfect score: 72

Sequence: 1 CGYWLTIWGC 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	63.9	218	1 Y232_SVNY3	Q55705 synchocyst
2	45	62.5	828	1 Q8ZBL4	Q8ZBL4 versinia pe
3	45	62.5	853	2 Q8DIA7	Q8DIA7 versinia pe
4	45	62.5	853	2 A860566	A860566 versinia pe
5	43.5	60.4	1993	2 P90670	P90670 aplysia cal
6	43	59.7	168	2 Q6H022	Q6H022 fremyella d
7	43	59.7	221	2 Q74JK6	Q74JK6 lactobacill
8	43	59.7	221	2 A8508923	A8508923 lactobaci
9	43	59.7	501	2 Q6C574	Q6C574 yarrowia li
10	43	59.7	646	2 Q8DKD6	Q8DKD6 synchococc
11	43	59.7	733	2 Q8H2N7	Q8H2N7 oryza sativ
12	43	59.7	1084	2 Q7WY20	Q7WY20 pseudomonas
13	42.5	59.0	152	2 Q96AC0	Q96AC0 homo sapien
14	42.5	59.0	159	2 Q6RW13	Q6RW13 homo sapien
15	42.5	59.0	159	2 Q9NRW9	Q9NRW9 homo sapien
16	42.5	59.0	159	2 Q86PL4	Q86PL4 homo sapien
17	42.5	59.0	159	2 A825556	A825556 homo sapi
18	42.5	59.0	263	2 Q7YIF1	Q7YIF1 cryptospori
19	42.5	59.0	1765	2 Q88457	Q88457 rattus norv
20	42	58.3	72	1 V635_BPML5	Q85245 mycobacteri
21	42	58.3	133	2 Q855L3	Q855L3 mycobacteri
22	42	58.3	284	2 Q6MC01	Q6MC01 parachlamy
23	42	58.3	284	2 C8F23898	C8F23898 parachlam
24	42	58.3	343	2 Q94FS2	Q94FS2 cajanus caj
25	42	58.3	389	2 Q83I51	Q83I51 tropheryma
26	42	58.3	441	2 Q92ZT6	Q92ZT6 rhizobium m
27	42	58.3	452	2 Q83G14	Q83G14 tropheryma
28	42	58.3	472	1 R8L_NITVU	Q95613 nitrobacter
29	42	58.3	473	2 Q9XD76	Q9XD76 nitrobacter
30	42	58.3	473	2 Q9XD77	Q9XD77 nitrobacter
31	42	58.3	473	2 Q8VQ84	Q8VQ84 nitrosospir

32	42	58.3	1308	2 Q8T6H2	Q8T6H2 dictyosteli
33	42	58.3	1498	1 MRP6_MOUSE	Q9RIE7 mus musculu
34	42	58.3	1502	1 MRP6_RAT	Q88369 rattus norv
35	42	58.3	1503	1 MRP6_HUMAN	Q95255 homo sapien
36	41.5	57.6	122	2 Q8N1Z5	Q8N1Z5 homo sapien
37	41.5	57.6	620	2 Q67544	Q67544 aquifex aeo
38	41	56.9	266	2 Q82Z85	Q82Z85 enterococcu
39	41	56.9	307	2 Q9M030	Q9M030 arabidopsis
40	41	56.9	376	2 Q6CBE4	Q6CBE4 yarrowia li
41	41	56.9	398	2 Q8IUD8	Q8IUD8 homo sapien
42	41	56.9	413	2 Q9HKA9	Q9HKA9 thermoplas
43	41	56.9	419	2 Q97VB7	Q97VB7 sulfolobus
44	41	56.9	432	2 Q6NUR0	Q6NUR0 homo sapien
45	41	56.9	432	2 A8H59368	A8H59368 homo sapi

ALIGNMENTS

RESULT 1

Y232_SVNY3
ID_Y232_SVNY3 STANDARD; PRT; 218 AA.
AC Q55705;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical protein slr0232.
GN OrderedLocNames=slr0232;
OS Synchocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Suglura M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Rep. 2:153-166(1995).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the deda family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; D64000; BAA10237.1; -;
CC F01; S76385; S76385.
CC InterPro; IPR000252; Deda.
CC Pfam; PF00597; Deda; 1.
CC Complete proteome; Hypothetical protein.
CC SEQUENCE 218 AA; 23781 MW; C04B5D7B7EA7F863 CRC64;

Query Match 63.9%; Score 46; DB 1; Length 218;

Best Local Similarity 66.7%; Pred. No. 50;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYWLTIWGC 9

|||||:

Db 73 CGYVWGRWG 81

RESULT 2

Q8ZBL4 PRELIMINARY; PRT; 828 AA.

ID Q8ZBL4

AC Q8ZBL4;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

```
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE ATP-dependent helicase.
GN Name=hrpB; OrderedLocusNames=YPO3394;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Tittball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL; AJ414156; CAC92624.1; -.
DR PIR; AD0412; AD0412.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR010225; DEAD_box_HrpB.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR007502; Helicase_dom.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF04408; HA2; 1.
DR PIRSF; PIRSF005496; ATP_hel_hrpB; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR TIGRFAMs; TIGR01970; DEAH_box_HrpB; 1.
KW ATP-binding; Complete proteome; Helicase; Hydrolase.
SQ SEQUENCE 828 AA; 91982 MW; 3A11F8883D5E583 CRC64;

Query Match 62.5%; Score 45; DB 2; Length 828;
Best Local Similarity 58.3%; Pred. No. 2.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Oy 1 CGYWLTI--WGC 10
Db 372 CGLWLELLSWG 383

RESULT 3
ID Q8DIA7 PRELIMINARY; PRT; 853 AA.
AC Q8DIA7; Q74XU2;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Helicase, ATP-dependent.
GN Name=hrpB; OrderedLocusNames=YPO291, y0794;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
```

```
RL J. Bacteriol. 184:4601-4611(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB013681; AM84381.1; -.
DR EMBL; AB017128; AAS60566.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0004386; F:helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR010225; DEAD_box_HrpB.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR007502; Helicase_dom.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF04408; HA2; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR PIRSF; PIRSF005496; ATP_hel_hrpB; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR TIGRFAMs; TIGR01970; DEAH_box_HrpB; 1.
KW ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 853 AA; 95004 MW; B3DB738A18665B42 CRC64;

Query Match 62.5%; Score 45; DB 2; Length 853;
Best Local Similarity 58.3%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Oy 1 CGYWLTI--WGC 10
Db 397 CGLWLELLSWG 408

RESULT 4
AAS60566 PRELIMINARY; PRT; 853 AA.
ID AAS60566
AC AAS60566;
DT 24-MAR-2004 (TReMBLrel. 27, Created)
DT 24-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 04-MAY-2004 (TReMBLrel. 27, Last annotation update)
DE Helicase, ATP-dependent.
GN HRPB OR YPO291.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017128; AAS60566.1; -.
DR EMBL; AB017128; AAS60566.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0004386; F:helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR010225; DEAD_box_HrpB.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR007502; Helicase_dom.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF04408; HA2; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR PIRSF; PIRSF005496; ATP_hel_hrpB; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR TIGRFAMs; TIGR01970; DEAH_box_HrpB; 1.
KW ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 853 AA; 95004 MW; B3DB738A18665B42 CRC64;

Query Match 62.5%; Score 45; DB 2; Length 853;
Best Local Similarity 58.3%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Oy 1 CGYWLTI--WGC 10
Db 397 CGLWLELLSWG 408
```



```
RESULT 5
P90670 PRELIMINARY; PRT; 1993 AA.
ID P90670;
AC P90670;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-MAY-2004 (TReMBLrel. 26, Last annotation update)
DE Sodium channel alpha-subunit SCAP1.
OS Aplysia californica (California sea hare).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Heterobranchia; Euthyneura; Opisthobranchia; Anaspidia;
OC Aplysioidea; Aplysiidae; Aplysia.
OX NCBI_TaxID=6500;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Nervous system;
RX MEDLINE=97238630; PubMed=9115644;
RA Dyer J.R., Johnston W.L., Castellucci V.F., Dunn R.J.;
RT "Cloning and tissue distribution of the Aplysia Na+ channel alpha-
RT subunit cDNA.";
RL DNA Cell Biol. 16:347-356 (1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sodium channel family.
DR EMBL; U66915; AAC47457.1; -
DR PIR; T30902; T30902.
DR HSSP; P04775; 1BYV.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0001518; C:voltage-gated sodium channel complex; IEA.
DR GO; GO:0005261; F:ion channel activity; IEA.
DR GO; GO:0005248; F:voltage-gated sodium channel activity; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001682; Ca/Na_pore.
DR InterPro; IPR002111; Cat_channel_TrpL.
DR InterPro; IPR005821; Ion_trans.
DR InterPro; IPR005820; M+channel_nlg.
DR InterPro; IPR001696; Na_channel.
DR InterPro; IPR010526; Na_trans_assoc.
DR Pfam; PF00520; Ion_trans_4.
DR Pfam; PF05512; Na_trans_assoc; 1.
DR PRINTS; PR00170; NACHANNEL.
KW Ion transport; Ionic channel; Sodium channel; Transmembrane;
KW Transport; Voltage-gated channel.
SQ SEQUENCE 1993 AA; 225896 MW; 33E174B9BF07E1A7 CRC64;

Query Match 60.4%; Score 43.5; DB 2; Length 1993;
Best Local Similarity 54.5%; Pred. No. 8.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 CGYML-TWGC 10
Db 942 CGEIESMWGC 952

RESULT 6
Q6H022 PRELIMINARY; PRT; 168 AA.
ID Q6H022;
AC Q6H022;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Fremyella diplosiphon (Calothrix PCC 7601).
OC Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Fremyella.
OX NCBI_TaxID=1197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FD33;
RA Stowe-Evans E.L., Ford J., Kehoe D.M.;
RT "Genomic DNA Microarray Analysis: Identification of New Genes
RT Regulated by Light Color in the Cyanobacterium Fremyella
RT diplosiphon.";
```

```
RL J. Bacteriol. 186:4338-4349 (2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FD33;
RA Stowe-Evans E., Ford J., Kehoe D.M.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY548455; AAT41947.1; -
KW Hypothetical protein.
SQ SEQUENCE 168 AA; 18932 MW; 1C9DB963D5210332 CRC64;

Query Match 59.7%; Score 43; DB 2; Length 168;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGYMLTWG 9
Db 122 CGYMLSLRG 130

RESULT 7
Q74JK6 PRELIMINARY; PRT; 221 AA.
ID Q74JK6;
AC Q74JK6;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hemolysin-like protein.
GN OrderedLocustNames=LJ1101;
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Fridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwaalen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollet B., Mercenier A., Klaenhammer T., Arigoni F., Schell M.A.;
RT "The genome sequence of the probiotic intestinal bacterium
RT Lactobacillus johnsonii NCC 533.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517 (2004).
DR EMBL; AE017203; AAS08923.1; -
DR InterPro; IPR004254; HlyIII_related.
DR InterPro; IPR005744; HlyIII.
DR Pfam; PF03006; HlyIII; 1.
DR TIGRFAMs; TIGR01065; hlyIII; 1.
KW Complete proteome.
SQ SEQUENCE 221 AA; 24721 MW; 1DBC78FF9810E152 CRC64;

Query Match 59.7%; Score 43; DB 2; Length 221;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYWLTVWG 9
Db 172 GFWLLVWG 179

RESULT 8
AAS08923 PRELIMINARY; PRT; 221 AA.
ID AAS08923;
AC AAS08923;
DT 02-MAR-2004 (TReMBLrel. 27, Created)
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)
DE Hemolysin-like protein.
GN LJ1101.
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]
```

```

RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Priddy R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwahlen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollet B., Mercenier A., Klammer T., Arigoni F., Schell M.A.;
RT "The genome sequence of the probiotic intestinal bacterium
RT Lactobacillus johnsonii NCC 533.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
DR EMBL; AE017203; AA089923.1; -.
SQ SEQUENCE 221 AA; 24721 MW; 1DBC78FF9810E152 CRC64;

Query Match 59.7%; Score 43; DB 2; Length 221;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GYWLTVWG 9
Db 172 GFWLLVWG 179

RESULT 9
-Q6C574
ID Q6C574 PRELIMINARY; PRT; 501 AA.
AC Q6C574
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Chromosome E of strain CLIB99 of Yarrowia lipolytica.
GN ORFNames=YALI0E20471g;
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG GENOLEVURES;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisarame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Deapons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikoleki M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekai F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpell C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382131; CAG79783.1; -.
SQ SEQUENCE 501 AA; 56411 MW; 91F08EFSFA63C60FB CRC64;

Query Match 59.7%; Score 43; DB 2; Length 501;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GYWLTVWG 10
Db 157 GYWLTVWG 165

RESULT 10
Q8DKD6
ID Q8DKD6 PRELIMINARY; PRT; 646 AA.
AC Q8DKD6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T110923 protein.
GN OrderedLocusNames=t110923;
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=22255144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Kato H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005372; BAC08475.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2CC; 1.
DR SMART; SM00331; PP2C_Sig; 1.
KW Complete proteome.
SQ SEQUENCE 646 AA; 71071 MW; 56C5A74F76652D56 CRC64;

Query Match 59.7%; Score 43; DB 2; Length 646;
Best Local Similarity 62.5%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 YWLTVWG 10
Db 415 YWINWGC 422

RESULT 11
Q8H2N7
ID Q8H2N7 PRELIMINARY; PRT; 733 AA.
AC Q8H2N7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein QJ1138_B05.118.
GN Names=QJ1138_B05.118;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005486; BAC16197.1; -.
DR Gramene; Q8H2N7; -.
DR InterPro; IPR007658; DUF594.
DR Pfam; PF04578; DUF594; 1.
KW Hypothetical protein.
SQ SEQUENCE 733 AA; 82939 MW; E95884DAD1DC2AC9 CRC64;

Query Match 59.7%; Score 43; DB 2; Length 733;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 YWLTVWG 9
Db 328 YWTVWG 334

```

```

RESULT 12
QWY20
ID Q7WY20 PRELIMINARY; PRT; 1084 AA.
AC Q7WY20;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Kcsc.
GN Name=rscs; ORFNames=RL038;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=PA14;
RC PubMed=14983043;
RA He J., Baldini R.L., Dezziel E., Saucier M., Zhang Q., Liberati N.T.,
RA Lee D., Urbach J., Goodman H.M., Rahme L.G.;
RT "The broad host range pathogen Pseudomonas aeruginosa strain PA14
RT carries two pathogenicity islands harboring plant and animal virulence
RT genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2530-2535(2004).
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
DR EMBL; AY273869; AAP84165.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component response regulator activity; IEA.
DR GO; GO:0000160; F:two-component sensor molecule activity; IEA.
DR GO; GO:0000160; F:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR003594; ATPbind ATPase.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kin_N.
DR InterPro; IPR008207; Hpt.
DR InterPro; IPR000014; PAS.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF00512; HisKA; 1.
DR Pfam; PF01627; Hpt; 1.
DR Pfam; PF00072; Response_reg; 1.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00388; HisKA; 1.
DR SMART; SM00091; PAS; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50894; HPT; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Kinase; Phosphorylation; Sensory transduction; Transferase.
SQ SEQUENCE 1084 AA; 119129 MW; C953FDD2F273BF1B CRC64;

Query Match 59.7%; Score 43; DB 2; Length 1084;
Best Local Similarity 66.7%; Pred. NO. 6e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GYWLTIWGC 10
DB 750 GAWLKAWGC 758

RESULT 13
Q96AC0
ID Q96AC0 PRELIMINARY; PRT; 152 AA.
AC Q96AC0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGTRAP protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Query Match 59.0%; Score 42.5; DB 2; Length 159;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

```

```

OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017328; AAH17328.1; -.
DR InterPro; IPR009436; AGTRAP.
DR Pfam; PF06396; AGTRAP; 1.
SQ SEQUENCE 152 AA; 16669 MW; 637C01214175C39 CRC64;

Query Match 59.0%; Score 42.5; DB 2; Length 152;
Best Local Similarity 70.0%; Pred. NO. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 2 GYW-LTIWGC 10
DB 14 GHWLLTTWGC 23

RESULT 14
Q6RW13
ID Q6RW13 PRELIMINARY; PRT; 159 AA.
AC Q6RW13;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Angiotensin II receptor-associated protein.
GN Name=AGTRAP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RA Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA Rajkumar N., Yi Q., Nickerson D.A.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY488088; AAR255556.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR009436; AGTRAP.
DR Pfam; PF06396; AGTRAP; 1.
KW Receptor.
SQ SEQUENCE 159 AA; 17419 MW; 7E1D5C7E79A56BC5 CRC64;

Query Match 59.0%; Score 42.5; DB 2; Length 159;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

```

QY 2 GYW-LTIWGC 10
|:| || |||
Db 14 GHMLLTWGC 23

RESULT 15

Q9NRW9 PRELIMINARY; PRT; 159 AA.
AC Q9NRW9 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ATRAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Ye R.D., He R.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF165187; AAF89547.1; -.
DR Genew; HGNC:13539; AGTRAP.
DR InterPro; IPR009436; AGTRAP.
DR Pfam; PF06396; AGTRAP; 1.
SQ SEQUENCE 159 AA; 17518 MW; 7E012C7E79AB6BC5 CRC64;

Query Match 59.0%; Score 42.5; DB 2; Length 159;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 2 GYW-LTIWGC 10
|:| || |||
Db 14 GHMLLTWGC 23

Search completed: December 30, 2004, 13:16:10
Job time : 94.2641 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:11:20 ; Search time 79.52 Seconds
(without alignments)
31.578 Million cell updates/sec

Title: US-10-046-922-67

Perfect score: 38

Sequence: 1 GYWXW 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	7	3 AAY76794	Somatostatin
2	38	100.0	7	5 ABP53964	VEGFR-3 b
3	38	100.0	7	5 ABP53418	Backbone
4	38	100.0	8	5 ABP53965	VEGFR-3 b
5	38	100.0	9	5 ABP53933	VEGFR-3 b
6	38	100.0	10	4 AAB99759	Rhesus D
7	38	100.0	10	5 ABP53931	VEGFR-3 b
8	38	100.0	10	5 ABP53968	VEGFR-3 b
9	38	100.0	10	5 ABP53932	VEGFR-3 b
10	38	100.0	11	2 AAW62148	Haemophil
11	38	100.0	11	8 ADJ25827	Tyrosine
12	38	100.0	12	4 AAB99769	Rhesus D
13	38	100.0	13	6 AAO26093	Fc region
14	38	100.0	13	8 ADJ50760	Human ser
15	38	100.0	16	5 AAU93268	Granulocyte
16	38	100.0	25	7 ADC99638	Cancer-re
17	38	100.0	29	4 ABB34417	Peptide #
18	38	100.0	29	4 AAM27898	Protein #
19	38	100.0	29	4 ABB19829	Human bra
20	38	100.0	29	4 AAM55207	Human bra
21	38	100.0	35	4 ABB41036	Peptide #
22	38	100.0	35	4 AAM34810	Protein #
23	38	100.0	35	4 ABB25111	Protein #
24	38	100.0	35	4 AAM74694	Human bon
25	38	100.0	35	4 AAM61892	Human bra

26	38	100.0	35	4 ABG56478	Human liv
27	38	100.0	35	5 ABG44505	Human pep
28	38	100.0	69	4 AAO13595	Human pol
29	38	100.0	74	4 AAU60558	Propionib
30	38	100.0	74	6 ABM57077	Propionib
31	38	100.0	91	4 AAU48276	Propionib
32	38	100.0	91	6 ABM44795	Propionib
33	38	100.0	100	3 AAB41381	Human ORF
34	38	100.0	116	2 AAR15437	Heavy cha
35	38	100.0	116	6 ABO27261	ICAM-1 bi
36	38	100.0	116	6 ABO27269	ICAM-1 bi
37	38	100.0	116	6 ABO27263	ICAM-1 bi
38	38	100.0	116	6 ABO27259	ICAM-1 bi
39	38	100.0	116	6 ABO27255	ICAM-1 bi
40	38	100.0	116	6 ABO27277	Humanised
41	38	100.0	116	6 ABO27273	Murine 1A
42	38	100.0	116	6 ABO27257	ICAM-1 bi
43	38	100.0	116	6 ABO27271	ICAM-1 bi
44	38	100.0	116	6 ABO27267	ICAM-1 bi
45	38	100.0	117	6 ABO27265	ICAM-1 bi

ALIGNMENTS

RESULT 1

AAY76794
ID AAY76794 standard; peptide; 7 AA.

XX AC AAY76794;

DT 20-APR-2000 (first entry)

DE Somatostatin analogue peptide 3181.

XX
KW Somatostatin analogue; therapy: cyclic peptide; autoimmune disease; endocrine disorder; cancer; diabetic-associated complication; diagnosis; gastrointestinal disorder; inflammatory disease; pancreatitis;
KW atherosclerosis; restenosis; post-surgical pain; VIP secretion inhibitor;
KW hormone-secreting tumour; hormone-dependent tumour; diarrhoea;
KW vasoactive intestinal peptide; non-insulin dependent diabetes mellitus.
XX
OS Synthetic.

XX Key Location/Qualifiers
FH Misc-difference 3 /note= "D-form residue"
FT Modified-site 7 /note= "Trp-NH2"

FT WO9965508-A1.

XX 23-DEC-1999.

XX 15-JUN-1999; 99WO-IL000329.

XX 19-JUN-1998; 98US-00100360.

PR 02-DEC-1998; 98US-00203389.

XX (PEPT-) PEPTOR LTD.

PI Hornik V, Afargan MW, Gellerman G;

DR WPI; 2000-136888/12.

XX Cyclic somatostatin analogs for inhibiting growth hormone secretion from anterior pituitary and as antiproliferative agents for the treatment of tumors.

XX Example 11; Page 61; 82pp; English.

XX This sequence represents a somatostatin analogue of the invention. The invention relates to a backbone cyclised somatostatin analogue that has

CC one building unit containing a nitrogen atom of the peptide backbone
 CC connected to a bridging group comprising an amide, thioether, thioester
 CC or disulphide. At least one building unit is connected via a bridging
 CC group to form a cyclic structure with a moiety selected from a second
 CC building unit, side chain of or N-terminal amino acid residue. A
 CC composition containing the analogue may be used for preventing disorders
 CC such as cancers, autoimmune diseases, endocrine disorders, diabetic-
 CC associated complications, gastrointestinal disorders, inflammatory
 CC diseases, pancreatitis, atherosclerosis, restenosis and post-surgical
 CC pain. It may also be used for diagnosing cancer. The backbone cyclic
 CC analogue is used for imaging the existence of metastases. Somatostatin
 CC analogues can be used for the treatment of patients with hormone-secreting
 CC and hormone-dependent tumours. They reduce diarrhoea through the
 CC inhibition of vasoactive intestinal peptide (VIP) secretion and by direct
 CC effect on intestinal secretion. Somatostatin analogues selective to type
 CC 2 and 5 receptors may be used for treatment of non-insulin dependent
 CC diabetes mellitus. They are useful for the prevention of atherosclerosis
 CC and restenosis. The analogues are metabolically stable, selective in
 CC their in-vivo activities and safe

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 3; Length 7;
 Best Local Similarity 57.1%; Pred. No. 1.7e+06;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
 |||:::
 Db 1 GYWKVCW 7

RESULT 2

ABP53964
 ID ABP53964 standard; peptide; 7 AA.

XX AC ABP53964;

XX DT 09-JAN-2003 (first entry)

XX DE VEGFR-3 binding peptide SEQ ID NO:67.

XX KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
 KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
 KW cytostatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
 KW vulnary; cell surface receptor; cancer; neovascularisation;
 KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
 KW diabetes; PDGF; platelet derived growth factor.

XX OS Homo sapiens.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 4..6
 FT /note= "X is any amino acid"

XX WO200257299-A2.

XX PD 25-JUL-2002.

XX PF 16-JAN-2002; 2002WO-IB000099.

XX PR 17-JAN-2001; 2001US-0262476P.

XX PA (LUDW-) LUDWIG INST CANCER RES.
 XX PA (LICN) LICENTIA LTD.

XX PI Alicalo K, Koivunen E, Kubo H;

XX WPI; 2002-691521/74.

XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
 PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
 PT such as cancer and diseases of neovascularization.

XX Claim 21; Page 81; 149pp; English.

XX The present invention describes an isolated peptide (I) that binds to and
 CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnary activities, and can be used in gene therapy.
 CC Compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
 CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of,
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 5; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
 |||:::
 Db 1 GYWXXXW 7

RESULT 3

ABP53418
 ID ABP53418 standard; peptide; 7 AA.

XX AC ABP53418;

XX DT 19-NOV-2002 (first entry)

XX DE Backbone cyclised somatostatin analogue PTR 3181.

XX KW Backbone cyclised somatostatin analogue; somatostatin; SRIF; analgesic;
 KW somatostatin release inhibiting factor; somatostatin receptor subtype;
 KW synthesis; antiarteriosclerotic; immunosuppressive; cytostatic; cancer;
 KW antidiabetic; antiinflammatory; somatostatin receptor ligand;
 KW atherosclerosis; autoimmune disease; diabetic-associated complication;
 KW endocrine disorder; inflammation; gastrointestinal disorder; restenosis;
 KW pancreatitis; post-surgical pain.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1
 FT /note= "N-terminally modified with Fmoc
 (fluorenylmethoxycarbonyl)"

FT Misc-difference 3

FT /note= "D form residue"

FT Modified-site 7
 FT /note= "amidated"

XX US2002052315-A1.

XX PD 02-MAY-2002.

XX PF 13-DEC-2000; 2000US-00734583.

XX PR 19-JUN-1998; 98US-00100360.

XX PR 02-DEC-1998; 98US-00203389.

XX PR 15-JUN-1999; 99WO-IL000329.

XX (HORN/) HORNIK V.

XX (AFAR/) AFARGAN M M.

XX (GELL/) GELLERMAN G.

XX PI Hornik V, Afargan MM, Gellerman G;

XX WPI; 2002-681319/73.
 DR New backbone cyclized somatostatin analogs are e.g. useful in the
 PT treatment of atherosclerosis, autoimmune diseases and cancers.
 XX
 PS Example 12; Page 21; 30pp; English.
 XX
 CC The present invention describes backbone cyclised somatostatin analogues
 CC (I) that incorporates at least one building unit containing one nitrogen
 CC atom of the peptide backbone connected to a bridging group (comprising an
 CC amide, thioether, thioester or disulfide) where at least one building
 CC unit is connected via the bridging group to form a cyclic structure with
 CC a moiety selected from the group consisting of a second building unit,
 CC the side chain of an amino acid residue of the sequence or the N-terminal
 CC amino acid residue. (I) has antiarteriosclerotic, immunosuppressive, and
 CC cytostatic, antidiabetic, antiinflammatory and analgesic activities, and
 CC can be used as a somatostatin receptor ligand. (I) are useful in the
 CC treatment of atherosclerosis, autoimmune diseases, cancers, diabetic-
 CC associated complications, endocrine disorders, inflammation,
 CC gastrointestinal disorders, pancreatitis, post-surgical pain, and
 CC restenosis. (I) can also be used in the diagnosis of cancer, by imaging
 CC the existence of metastases, it being labeled with a detectable probe.
 CC The present sequence represents a backbone cyclised somatostatin analogue
 CC from the present invention
 XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 38; DB 5; Length 7;
 Best Local Similarity 57.1%; Pred. No. 1.7e+06;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GYVXXXW 7
 Db |||::||
 1 GYVRCVW 7
 RESULT 4
 ABP53965
 ID ABP53965 standard; peptide; 8 AA.
 XX
 AC ABP53965;
 XX
 DT 09-JAN-2003 (first entry)
 XX
 DE VEGFR-3 binding peptide SEQ ID NO:68.
 XX
 KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
 KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
 KW cytotatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
 KW vulnary; cell surface receptor; cancer; neovascularisation;
 KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
 KW diabetes; PDGF; platelet derived growth factor.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 4..6
 FT /note= "X is any amino acid"
 FT Misc-difference 8
 FT /note= "any amino acid"
 XX
 PN WO200257299-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 16-JAN-2002; 2002WO-IB000099.
 XX
 PR 17-JAN-2001; 2001US-0262476P.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 PA (LICN) LICENTIA LTD.
 XX
 PI Alitálo K, Koivunen E, Kubo H;
 XX
 DR WPI; 2002-691521/74.
 XX
 PT New isolated peptide that inhibits VEGF-C and VEGF-D, useful for

XX Alitálo K, Koivunen E, Kubo H;
 XX WPI; 2002-691521/74.
 DR
 XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
 PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
 PT such as cancer and diseases of neovascularization.
 XX
 PS Claim 22; Page 81; 149pp; English.
 XX
 CC The present invention describes an isolated peptide (I) that binds to and
 CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnary activities, and can be used in gene therapy.
 CC Compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
 CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 XX
 SQ Sequence 8 AA;
 Query Match 100.0%; Score 38; DB 5; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GYVXXXW 7
 Db |||::|||
 1 GYVXXXW 7
 RESULT 5
 ABP53933
 ID ABP53933 standard; peptide; 9 AA.
 XX
 AC ABP53933;
 XX
 DT 09-JAN-2003 (first entry)
 XX
 DE VEGFR-3 binding peptide SEQ ID NO:36.
 XX
 KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
 KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
 KW cytotatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
 KW vulnary; cell surface receptor; cancer; neovascularisation;
 KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
 KW diabetes; PDGF; platelet derived growth factor.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200257299-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 16-JAN-2002; 2002WO-IB000099.
 XX
 PR 17-JAN-2001; 2001US-0262476P.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 PA (LICN) LICENTIA LTD.
 XX
 PI Alitálo K, Koivunen E, Kubo H;
 XX
 DR WPI; 2002-691521/74.
 XX
 PT New isolated peptide that inhibits VEGF-C and VEGF-D, useful for

```

PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
PT such as cancer and diseases of neovascularization.
XX
PS Claim 14; Page 80; 149pp; English.
XX
CC The present invention describes an isolated peptide (I) that binds to and
CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
CC antidiabetic and vulnerary activities, and can be used in gene therapy.
CC Compositions and methods from the present invention are useful for
CC diagnosing, evaluating and treating disorders mediated by the activity of
CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
CC liver, spleen, kidney, lymph node, small intestine, blood cells,
CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
CC chronic hepatitis, haemangiomas and diabetes. The present sequence
CC represents a specifically claimed VEGFR-3 binding peptide from the
CC present invention
XX
XX Sequence 9 AA;
SQ
Query Match 100.0%; Score 38; DB 5; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.7e+06;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYWXIXW 7
DB 2 GYWDTW 8
RESULT 6
AAB99759
ID AAB99759 standard; peptide; 10 AA.
AC AAB99759;
XX
DT 21-SEP-2001 (first entry)
DE
XX Rhesus D antibody binding peptide SEQ ID NO:4.
XX Rhesus D antibody binding peptide; Rhesus D; Rhd; identification;
KW anti-Rhesus D antibody; immunogen; epitope; diagnosis; therapy;
KW prophylaxis; haemolytic disease of the newborn; HDN; ITP;
KW idiopathic thrombocytopenic purpura; immunoglobulin.
XX
OS Homo sapiens.
OS Synthetic.
PN EP1106625-A1.
XX
PD 13-JUN-2001.
XX
PF 17-NOV-1999; 99EP-00122858.
XX
PR 17-NOV-1999; 99EP-00122858.
XX
PA (ZLBB-) ZLB BIOPLASMA AG.
XX
XX Miescher S, Hofmann A, Fisch I;
XX WPI; 2001-383568/41.
XX
XX Novel peptides capable of binding Rhesus D antibodies are used to
PT manufacture an agent for the diagnosis, therapy or prophylaxis of
PT diseases associated with Rhesus D antigen, e.g. hemolytic disease of the
PT newborn (HDN).
XX
XX Claim 1; Page 12; 19pp; English.
XX
CC The present sequence represents a peptide capable of binding Rhesus D
CC antibodies (I). Also described in the present invention are: (1) a
CC nucleic acid (II) encoding (I); (2) a vector (III) comprising one or
CC (II) operably linked to an expression control system; (3) a cell (IV)

```

```

CC comprising (II) or (III); (4) preparing (I); (5) identifying (M1)
CC peptides having immunologic properties of Rhesus D protein epitopes
CC comprising subjecting an antibody/antibody fragment recognising an
CC epitope of Rhesus D protein to several panning rounds with a phage
CC display library, and identifying immunogenic peptide sequences which are
CC mimotopes which differ in their amino acid sequence from the amino acid
CC sequences of Rhesus D protein; and (6) peptides (V) with immunological
CC properties of Rhesus D protein epitopes obtained by (M1). (I) is used to
CC manufacture an agent for the diagnosis, therapy or prophylaxis of
CC diseases associated with Rhesus D antigen, e.g. haemolytic disease of the
CC newborn (HDN) or idiopathic thrombocytopenic purpura (ITP), for the
CC manufacture of an affinity reagent for anti-Rhesus D antibodies purified
CC or removed from body fluids or immunoglobulin preparations. Using (I) as
CC an immunogen to raise anti-Rhesus D antibodies avoids using immunisation
CC with foreign erythrocytes thereby avoiding the risk of transmission of
CC viral diseases like AIDS and hepatitis B
XX
XX Sequence 10 AA;
SQ
Query Match 100.0%; Score 38; DB 4; Length 10;
Best Local Similarity 57.1%; Pred. No. 44;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYWXIXW 7
DB 1 GYWSAKW 7
RESULT 7
ABP53931
ID ABP53931 standard; peptide; 10 AA.
XX
AC ABP53931;
XX
DT 09-JAN-2003 (first entry)
DE
XX VEGFR-3 binding peptide SEQ ID NO:34.
XX Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
KW cytostatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
KW vulnerary; cell surface receptor; cancer; neovascularisation;
KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
KW diabetes; PDGF; platelet derived growth factor.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 1 /note= "any amino acid"
FT Misc-difference 10 /note= "any amino acid"
XX
XX WO200257299-A2.
XX
XX 25-JUL-2002.
XX
XX 16-JAN-2002; 2002WO-IB0000099.
XX
XX 17-JAN-2001; 2001US-0262476P.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX (LICN) LICENTIA LTD.
XX
XX Alitalo K, Koivunen E, Kubo H;
XX WPI; 2002-691521/74.
XX
XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
PT such as cancer and diseases of neovascularization.
XX

```


PS Claim 12; Page 80; 149pp; English.

XX The present invention describes an isolated peptide (I) that binds to and

CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)

CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,

CC antidiabetic and vulnerary activities, and can be used in gene therapy.

CC Compositions and methods from the present invention are useful for

CC diagnosing, evaluating and treating disorders mediated by the activity of

CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,

CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,

CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of

CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,

CC chronic hepatitis, haemangiomas and diabetes. The present sequence

CC represents a specifically claimed VEGFR-3 binding peptide from the

CC present invention

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 38; DB 5; Length 10;

Best Local Similarity 57.1%; Pred. No. 44;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7

DB 2 GYWLTIW 8

RESULT 8

ABP53968

ID ABP53968 standard; peptide; 10 AA.

XX AC ABP53968;

XX 09-JAN-2003 (first entry)

XX VEGFR-3 binding peptide SEQ ID NO:73.

XX Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;

KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;

KW cytostatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;

KW vulnerary; cell surface receptor; cancer; neovascularisation;

KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;

KW diabetes; PDGF; platelet derived growth factor.

XX Homo sapiens.

OS Synthetic.

OS Key Location/Qualifiers

FH Misc-difference 5...7

FT Misc-difference 9 /note= "X is any amino acid"

FT Misc-difference 9 /note= "X is any amino acid"

XX WO200257299-A2.

XX 25-JUL-2002.

XX 16-JAN-2002; 2002WO-IB0000099.

XX 17-JAN-2001; 2001US-0262476P.

XX (LUDW-) LUDWIG INST CANCER RES.

PA (LICN) LICENTIA LTD.

XX Alitalo K, Koivunen E, Kubo H;

XX WPI; 2002-691521/74.

XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for

PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,

PT such as cancer and diseases of neovascularization.

XX

PS Claim 13; Page 80; 149pp; English.

XX The present invention describes an isolated peptide (I) that binds to and

CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)

CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,

CC antidiabetic and vulnerary activities, and can be used in gene therapy.

CC Compositions and methods from the present invention are useful for

CC diagnosing, evaluating and treating disorders mediated by the activity of

CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,

CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,

CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of

CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,

CC chronic hepatitis, haemangiomas and diabetes. The present sequence

CC represents a VEGFR-3 binding peptide, which is given in the

CC exemplification of the present invention

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 38; DB 5; Length 10;

Best Local Similarity 100.0%; Pred. No. 44;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7

DB 2 GYWXXXW 8

RESULT 9

ABP53932

ID ABP53932 standard; peptide; 10 AA.

XX AC ABP53932;

XX 09-JAN-2003 (first entry)

XX VEGFR-3 binding peptide SEQ ID NO:35.

XX Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;

KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;

KW cytostatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;

KW vulnerary; cell surface receptor; cancer; neovascularisation;

KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;

KW diabetes; PDGF; platelet derived growth factor.

XX Homo sapiens.

OS Synthetic.

OS WO200257299-A2.

XX 25-JUL-2002.

XX 16-JAN-2002; 2002WO-IB0000099.

XX 17-JAN-2001; 2001US-0262476P.

XX (LUDW-) LUDWIG INST CANCER RES.

PA (LICN) LICENTIA LTD.

XX Alitalo K, Koivunen E, Kubo H;

XX WPI; 2002-691521/74.

XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for

PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,

PT such as cancer and diseases of neovascularization.

XX

PS Claim 13; Page 80; 149pp; English.

XX The present invention describes an isolated peptide (I) that binds to and

CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)

CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,

CC antidiabetic and vulnerary activities, and can be used in gene therapy.

CC Compositions and methods from the present invention are useful for

CC diagnosing, evaluating and treating disorders mediated by the activity of

CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,

CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,

CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of

CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,

CC chronic hepatitis, haemangiomas and diabetes. The present sequence

CC represents a VEGFR-3 binding peptide, which is given in the

CC exemplification of the present invention

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 38; DB 5; Length 10;

Best Local Similarity 100.0%; Pred. No. 44;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7

DB 2 GYWXXXW 8

RESULT 9

ABP53932

ID ABP53932 standard; peptide; 10 AA.

XX AC ABP53932;

XX 09-JAN-2003 (first entry)

XX VEGFR-3 binding peptide SEQ ID NO:35.

XX Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;

KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;

KW cytostatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;

KW vulnerary; cell surface receptor; cancer; neovascularisation;

KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;

KW diabetes; PDGF; platelet derived growth factor.

XX Homo sapiens.

OS Synthetic.

OS WO200257299-A2.

XX 25-JUL-2002.

XX 16-JAN-2002; 2002WO-IB0000099.

XX 17-JAN-2001; 2001US-0262476P.

XX (LUDW-) LUDWIG INST CANCER RES.

PA (LICN) LICENTIA LTD.

XX Alitalo K, Koivunen E, Kubo H;

XX WPI; 2002-691521/74.

XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for

PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,

PT such as cancer and diseases of neovascularization.

XX

CC Compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
 CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 XX
 SQ Sequence 10 AA;

Query Match 100.0%; Score 38; DB 5; Length 10;

Best Local Similarity 57.1%; Pred. No. 44;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
 |||::|
 Db 2 GYWLTIW 8

RESULT 10

ID AAW62148 standard; peptide; 11 AA.

XX AAW62148;

XX 16-SEP-1998 (first entry)

DE Haemophilus influenzae tyrosine tRNA synthetase binding peptide 21.

KW Identification; ligand; biological activity; target-binding;

KW drug screening; library; inhibitory ligand.

XX Synthetic.

OS Haemophilus influenzae.

XX WO9819162-A1.

XX 07-MAY-1998.

PF 31-OCT-1997; 97WO-US019638.

PR 31-OCT-1996; 96US-00740671.

XX (NOVA-) NOVALON PHARM CORP.

PI Fowlkes DM, Kay BK, Frelinger JA, Hyde-Deruysscher RP;

DR WPI; 1998-272389/24.

XX Identifying ligands which mediate biological activity of a protein - by
 PT identifying target-binding ligands and screening a library for ligands
 PT which inhibit target-binding ligand mediated activity.

PS Example 5; Page 100; 143pp; English.

XX A method has been developed for identifying a ligand which mediates the
 CC biological activity of a target protein (T) by inhibiting the binding of
 CC (T) to a binding partner. The method comprises: (a) screening a first
 CC combinatorial library comprising first member ligands for binding to the
 CC target-binding ligands (TBLs), to identifying one or more TBLs; (b)
 CC screening a second library comprising second member ligands for the
 CC ability to inhibit the binding of one or more of the TBLs to the target
 CC protein, and so obtaining one or more inhibitory ligands; and (c)
 CC determining which of the inhibitory ligands can mediate a biological
 CC activity of the target protein. The present sequence represents a
 CC potential binding peptide for Haemophilus influenzae tyrosine tRNA
 CC synthetase from an example of the present invention. The method can be
 CC used for identifying drugs which can mediate the biological activity of a
 CC target protein. It can be used to identify the biological activity of a
 CC target protein whose biological function is not known and perhaps cannot

CC be determined directly. The method can also be used to identify new
 CC inhibitory ligands of specific target proteins. The method provides high
 CC throughput screens which are essentially identical for similar and
 CC dissimilar targets, bypassing the need to develop distinct assays for
 CC biochemically diverse targets
 XX
 SQ Sequence 11 AA;

Query Match 100.0%; Score 38; DB 2; Length 11;

Best Local Similarity 57.1%; Pred. No. 48;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
 |||::|
 Db 1 GIWWDPM 7

RESULT 11

ID ADJ25827 standard; peptide; 11 AA.

XX ADJ25827;

XX 20-MAY-2004 (first entry)

DE Tyrosine tRNA synthetase binding peptide #21.

KW ligand identification; peptide library;

KW complementary combinatorial library; tyrosine tRNA synthetase.

OS Synthetic.

XX US6617114-B1.

XX 09-SEP-2003.

PF 30-APR-1998; 98US-00069827.

XX 31-OCT-1996; 96US-00740671.

PR 31-OCT-1997; 97WO-US019638.

PR 31-MAR-1998; 98US-00050359.

XX (KARO-) KARO BIO AB.

PI Fowlkes DM, Kay BK, Frelinger JA, Hyde-Deruysscher RP;

XX WPI; 2004-068186/07.

XX Identification of ligand that can mediate biological activity of target
 PT protein, comprises screening first combinatorial library having first
 PT member ligands for binding to target protein to identify target-binding
 PT ligand(s).

PS Example 5; SEQ ID NO 87; 98pp; English.

XX The invention relates to a method of identifying a ligand that can
 CC mediate the biological activity of target protein via inhibition of the
 CC binding of target protein to a binding partner ligand comprising
 CC screening first combinatorial library having first member ligands for
 CC binding to target protein to identify target-binding ligand(s). The
 CC method is useful for identifying ligands that can mediate the biological
 CC activity of target proteins via inhibition of the binding of target
 CC protein to a binding partner ligand. The invention does not require that
 CC the natural binding partner be used as reagent. The need for the natural
 CC binding partner is obviated with the use of complementary combinatorial
 CC libraries. The present sequence is used in the exemplification of the
 CC present invention.

XX Sequence 11 AA;

Query Match 100.0%; Score 38; DB 8; Length 11;

Best Local Similarity 57.1%; Pred. No. 48;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
 DB 1 GYWPDW 7

RESULT 12
 ID AAB99769 standard; peptide; 12 AA.
 AC AAB99769;
 DT 21-SEP-2001 (first entry)
 DE Rhesus D antibody related peptide #5.
 XX Rhesus D antibody binding peptide; Rhesus D; Rhd; identification;
 KW anti-Rhesus D antibody; immunogen; epitope; diagnosis; therapy;
 KW prophylaxis; haemolytic disease of the newborn; HDN; ITP; cyclic;
 KW idiopathic thrombocytopenic purpura; immunoglobulin; circular.
 XX Homo sapiens.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Disulfide-bond 1..12
 FT
 XX EP1106625-A1.
 PN 13-JUN-2001.
 PD 17-NOV-1999; 99EP-00122858.
 PF 17-NOV-1999; 99EP-00122858.
 XX (ZLBB-) ZLB BIOPLASMA AG.
 PA Miescher S, Hofmann A, Fisch I;
 PI WPI; 2001-383568/41.
 DR Novel peptides capable of binding Rhesus D antibodies are used to
 PT manufacture an agent for the diagnosis, therapy or prophylaxis of
 PT diseases associated with Rhesus D antigen, e.g. hemolytic disease of the
 PT newborn (HDN).
 XX Example 1; Page 8; 19pp; English.
 XX The present invention describes peptides capable of binding Rhesus D
 CC antibodies (I). Also described in the present invention are: (1) a
 CC nucleic acid (II) encoding (I); (2) a vector (III) comprising one or more
 CC (II) operably linked to an expression control system; (3) a cell (IV)
 CC comprising (II) or (III); (4) preparing (I); (5) identifying (MI)
 CC peptides having immunologic properties of Rhesus D protein epitopes
 CC comprising subjecting an antibody/antibody fragment recognising an
 CC epitope of Rhesus D protein to several panning rounds with a phage
 CC display library, and identifying immunogenic peptide sequences which are
 CC mimotopes which differ in their amino acid sequence from the amino acid
 CC sequences of Rhesus D protein; and (6) peptides (V) with immunological
 CC properties of Rhesus D protein epitopes obtained by (MI). (I) is used to
 CC manufacture an agent for the diagnosis, therapy or prophylaxis of
 CC diseases associated with Rhesus D antigen, e.g. haemolytic disease of the
 CC newborn (HDN) or idiopathic thrombocytopenic purpura (ITP), for the
 CC manufacture of an affinity reagent for anti-Rhesus D antibodies purified
 CC or removed from body fluids or immunoglobulin preparations. Using (I) as
 CC an immunogen to raise anti-Rhesus D antibodies avoids using immunisation
 CC with foreign erythrocytes thereby avoiding the risk of transmission of
 CC viral diseases like AIDS and hepatitis B. The present sequence represents
 CC an anti-Rhesus D (Rhd) antibody related peptide which is used in an
 CC example from the present invention
 XX Sequence 12 AA;

Query Match 100.0%; Score 38; DB 4; Length 12;
 Best Local Similarity 57.1%; Pred. No. 53;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
 DB 2 GYWSAKW 8

RESULT 13
 ID AAO26093 standard; peptide; 13 AA.
 AC AAO26093;
 DT 03-APR-2003 (first entry)
 XX Fc region binding peptide SEQ ID No 73.
 KW Immunoglobulin Fc region; binding; whole blood; plasma; transgenic milk;
 KW antibody response; half-life; stability; circulatory system.
 XX Unidentified.
 OS WO200286070-A2.
 XX 31-OCT-2002.
 PD 18-APR-2002; 2002WO-US012492.
 PF 18-APR-2001; 2001US-0284534P.
 PR (DYAX-) DYAX CORP.
 PA Rondon IJ, Wu Q, Ley AC, Stochl M, Ransohoff TC, Potter MD;
 PI WPI; 2003-201220/19.
 DR New polypeptides, useful as binding molecules for detecting, isolating or
 PT purifying immunoglobulin Fc-region polypeptides present in a solution, or
 PT for regulating or preventing an antibody response.
 XX Claim 3; Page 76; 152pp; English.
 XX The invention relates to novel isolated polypeptides comprising a
 CC sequence that binds an immunoglobulin Fc region. The polypeptides are
 CC useful as binding molecules for detecting, isolating or purifying
 CC immunoglobulin Fc-region polypeptides present in a solution, e.g. whole
 CC blood, plasma or transgenic milk. The Fc-region binding polypeptides are
 CC also useful for regulating or preventing an antibody response, or for
 CC increasing the half-life and over all stability of a therapeutic or
 CC diagnostic compound that is administered to or enters the circulatory
 CC system of an individual. This sequence represents an Fc region binding
 CC peptide of the invention
 XX Sequence 13 AA;

Query Match 100.0%; Score 38; DB 6; Length 13;
 Best Local Similarity 57.1%; Pred. No. 57;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
 DB 1 GYWCNVW 7

RESULT 14
 ID ADJ50760 standard; peptide; 13 AA.
 AC ADJ50760;
 XX 06-MAY-2004 (first entry)

```

XX DE Human serum albumin binding peptide, Seq ID No 297.
XX KW human serum albumin; HSA; serum; blood; tumour; human.
XX OS Homo sapiens.
XX PN WO2003106493-A1.
XX PD 24-DEC-2003.
XX PF 16-JUN-2003; 2003WO-US018896.
XX PR 14-JUN-2002; 2002US-0388642P.
XX PA (DYAX-) DYAX CORP.
XX PI Sato AK, Dawson BM;
XX PS WPI; 2004-082161/08.
XX PT Evaluating sample comprising soluble serum protein by forming complex
XX PT comprising serum protein and physically associated compounds using
XX PT peptide ligand that specifically binds with proteins, which is separated
XX PT and evaluated.
XX PS Disclosure; SEQ ID NO 297; 191pp; English.
XX CC The invention relates to a method of evaluating sample by providing a
XX CC soluble serum protein (I), one or more compounds physically associated
XX CC with (I), and a (I)-binding agent that comprises a peptide that
XX CC specifically binds to (I), allowing the (I)-binding agent to bind to (I)
XX CC to form a complex including one or more compounds physically associated
XX CC with (I), separating the complex from one or more components of the
XX CC sample, and evaluating one or more of the physically associated
XX CC compounds. The sample comprises blood or serum, or is obtained from a
XX CC biopsy. The sample may also be obtained from a tumour or a region within
XX CC 5 mm of a tumour. The method is useful for detecting modulators that
XX CC modulate interaction of serum protein-binding compound and serum protein
XX CC and for identifying binding ligands for serum protein. The present
XX CC sequence represents a serum albumin-binding peptide identified using the
XX CC method of the invention.
XX SQ Sequence 13 AA;
XX Query Match 100.0%; Score 38; DB 8; Length 13;
XX Best Local Similarity 57.1%; Pred. No. 57;
XX Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
OY 1 GYWXXW 7
DB 1 GYWCNVW 7
|||:::|
|||:::|

RESULT 15
AAU93268
ID AAU93268 standard; peptide; 16 AA.
XX AC AAU93268;
XX DT 02-JUL-2002 (first entry)
XX DE Granulocyte-colony stimulating factor receptor binding peptide #74.
XX KW G-CSFR; granulocyte-colony stimulating factor receptor; cytokine;
XX KW haematopoietic growth factor; neutrophil proliferation; AIDS;
XX KW neutrophil differentiation; acquired immunodeficiency syndrome;
XX KW chemotherapy-induced neutropaenia; community acquired pneumonia;
XX KW depressed neutrophil count; immunostimulant.
XX OS Synthetic.
XX PN WO200207676-A2.

XX PD 31-JAN-2002.
XX PF 20-JUL-2001; 2001WO-US023046.
XX PR 20-JUL-2000; 2000US-00620091.
XX PA (GLAX ) GLAXO GROUP LTD.
XX PI Cwirla SE, Balu P, Duffin DJ, Piplani S, Meeowen-Merrill B;
XX PI Schatz PJ;
XX PS WPI; 2002-329382/36.
XX PT Novel compounds, useful for treating depressed neutrophil count, comprise
XX PT peptide chains of approximately 6 to 40 amino acids in length that bind
XX PT to granulocyte-colony stimulating factor receptor.
XX PS Claim 4; Page 53; 90pp; English.
XX CC The invention relates to compounds comprising a peptide chain
XX CC approximately 6 to 40 amino acids in length that binds to granulocyte-
XX CC colony stimulating factor receptor (G-CSFR). The compounds contain
XX CC specific sequences of the generic peptides appearing as AAU79402-AAU79406
XX CC and the generic sequences XV 1XV 2XV 3XV 4XV 5XV 6XV 7XV 8 (where XV 1 =
XX CC E, C, Q, V or Y; XV 2 = E, A, L, M, S, W or Q; XV 3 = K, R or T; XV 4 =
XX CC L, A or V; XV 5 = R, A, M, H, E, V, L, G, D, Q or S; XV 6 = E or V; XV 7
XX CC = A or G; and XV 8 = R, H, G or L) and XVI 1XVI 2XVI 3XVI 4XVI 5
XX CC EXVI 6XVI 7XVI 8XVI 9 (where XVI 1 = A, E or G; XVI 2 = E, H or D; XVI 3
XX CC = R or G; XVI 4 = K, Y, M, N, Q, R, D, I, S or E; XVI 5 = A, S or P;
XX CC XVI 6 = E, D, T, Q, K or R; XVI 7 = R, W, K, L, S, A or Q; XVI 8 = R or E
XX CC ; and XVI 9 = W, G or R). The compounds are used for treating conditions
XX CC associated with depressed neutrophil count e.g. chemotherapy- induced
XX CC neutropaenia, AIDS-induced neutropaenia or community-acquired pneumonia-
XX CC induced pneumonia. The compounds are useful as in vitro as tools for
XX CC understanding the biological role of granulocyte-colony stimulating
XX CC factor (G-CSF) a haematopoietic growth factor and cytokine that stimulates
XX CC neutrophil proliferation and differentiation), including evaluation of
XX CC many factors thought to influence, and be influenced by, production of
XX CC white blood cells, in the development of compounds that bind to G-CSFR,
XX CC as reagents for detecting G-CSF receptor or related receptor on living
XX CC cells, fixed cells, in biological fluid, in tissue homogenates or in
XX CC purified natural biological materials, in situ staining, fluorescence-
XX CC activated cell sorting (FACS), Western blotting or enzyme-linked
XX CC immunoadsorbent assay (ELISA), in receptor purification or in purifying
XX CC cells expressing G-CSFR on the cell surface (or inside permeabilised
XX CC cells) as a commercial research reagent for various medical and
XX CC diagnostic uses or to treat a disease that would benefit from the ability
XX CC to of a compound to mimic the effects of G-CSF in vivo. The compounds
XX CC bind specifically to G-CSFR and allow for studies of biological
XX CC activities mediated by the receptor and for the treatment of diseases,
XX CC disorders and conditions that would benefit from activating or
XX CC inactivating G-CSFR. The present sequence is a G-CSFR binding peptide of
XX CC the invention
XX SQ Sequence 16 AA;
XX Query Match 100.0%; Score 38; DB 5; Length 16;
XX Best Local Similarity 57.1%; Pred. No. 70;
XX Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
OY 1 GYWXXW 7
DB 1 GYWCDFW 7
|||:::|
|||:::|

Search completed: January 3, 2005, 16:27:22
Job time : 81.52 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:19:47 ; Search time 20.16 Seconds
(without alignments)
23.027 Million cell updates/sec

Title: US-10-046-922-67

Perfect score: 38

Sequence: 1 GYWXW 7

Scoring table: <BLZSUM62DX>

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	11	4	US-09-069-827A-87
2	38	100.0	16	4	US-09-620-091-81
3	38	100.0	177	4	US-09-543-681A-7620
4	38	100.0	252	3	US-09-502-653-10
5	38	100.0	263	4	US-09-610-905-12
6	38	100.0	274	4	US-09-248-796A-15791
7	38	100.0	278	4	US-09-145-828A-11
8	38	100.0	278	4	US-09-903-456-18
9	38	100.0	339	4	US-09-252-991A-26841
10	38	100.0	342	4	US-09-252-991A-32031
11	38	100.0	362	1	US-08-415-751-6
12	38	100.0	367	4	US-09-248-796A-15188
13	38	100.0	478	4	US-09-107-532A-6868
14	38	100.0	492	4	US-09-107-532A-6945
15	38	100.0	499	4	US-09-252-991A-23328
16	38	100.0	500	4	US-09-252-991A-21214
17	38	100.0	543	4	US-09-252-991A-18697
18	38	100.0	865	4	US-09-252-991A-19339
19	38	100.0	1498	4	US-09-792-616-9
20	38	100.0	1503	4	US-09-792-616-3
21	34	89.5	34	4	US-09-270-767-60715
22	34	89.5	43	2	US-08-488-161-69
23	34	89.5	43	3	US-09-273-685-69
24	34	89.5	43	5	PCT-US95-11934-69
25	34	89.5	55	4	US-09-621-976-7633
26	34	89.5	101	3	US-09-374-135-4
27	34	89.5	158	4	US-09-270-767-35277

28 34 89.5 158 4 US-09-270-767-50494 Sequence 50494, A
29 34 89.5 197 3 US-09-112-248-2 Sequence 2, Appli
30 34 89.5 227 4 US-09-489-039A-8752 Sequence 8752, Ap
31 34 89.5 233 4 US-09-248-796A-20388 Sequence 20388, A
32 34 89.5 255 4 US-09-270-767-45223 Sequence 45223, A
33 34 89.5 328 4 US-09-452-937A-34 Sequence 34, Appl
34 34 89.5 417 4 US-09-248-796A-18939 Sequence 18939, A
35 34 89.5 458 4 US-09-543-681A-6324 Sequence 6324, A
36 34 89.5 519 4 US-09-198-452A-479 Sequence 479, App
37 34 89.5 580 4 US-09-270-767-43086 Sequence 43086, A
38 34 89.5 602 4 US-09-248-796A-16589 Sequence 16589, A
39 34 89.5 625 4 US-09-252-991A-28537 Sequence 28537, A
40 34 89.5 639 4 US-09-270-767-61294 Sequence 61294, A
41 34 89.5 764 4 US-09-270-767-45772 Sequence 45772, A
42 33 86.8 13 4 US-09-069-827A-86 Sequence 86, Appl
43 33 86.8 19 4 US-09-794-529B-8 Sequence 8, Appl
44 33 86.8 19 4 US-09-794-517A-8 Sequence 8, Appl
45 33 86.8 19 4 US-09-011-645B-8 Sequence 8, Appl

ALIGNMENTS

RESULT 1

US-09-069-827A-87

; Sequence 87, Application US/09069827A

; Patent No. 6617114

; GENERAL INFORMATION:

APPLICANT: FOWLKES, Dana M

KAY, Brian K

FRELINGER, Jeffrey A

HYDE-DERUYSCHE, Robin P

TITLE OF INVENTION: IDENTIFICATION OF DRUGS USING

NUMBER OF SEQUENCES: 178

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.

STREET: 624 Ninth Street N.W., Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLYING APPLICATION NUMBER: US/09/069,827A

FILING DATE: 30-Apr-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/050,359

FILING DATE: 31-MAR-1998

APPLICATION NUMBER: PCT/US97/19638

FILING DATE: 31-OCT-1997

APPLICATION NUMBER: US 08/740,671

FILING DATE: 31-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: COOPER, Iver P

REGISTRATION NUMBER: 28,005

REFERENCE/DOCKET NUMBER: FOWLKES=4C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-5197

TELEFAX: (202) 737-3528

INFORMATION FOR SEQ ID NO: 87:

SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 87:

```
US-09-069-827A-87
Query Match      100.0%; Score 38; DB 4; Length 11;
Best Local Similarity 57.1%; Pred. No. 22;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 1 GYWPW 7

RESULT 2
US-09-620-091-81
; Sequence 81, Application US/09620091
; Patent No. 6716811
; GENERAL INFORMATION:
; APPLICANT: CWRILA, STEVEN E.
; APPLICANT: BALU, PALANI
; APPLICANT: DUFFIN, DAVID J.
; APPLICANT: PIPLANI, SUNILA
; APPLICANT: MERRILL, BARBARA MCEOWEN
; APPLICANT: SCHATZ, PETER JOSEPH
; TITLE OF INVENTION: COMPOUNDS HAVING AFFINITY FOR THE GRANULOCYTE-COLONY
; TITLE OF INVENTION: STIMULATING FACTOR RECEPTOR (G-CSFR) AND ASSOCIATED
; TITLE OF INVENTION: USES
; FILE REFERENCE: 0300-0014
; CURRENT APPLICATION NUMBER: US/09/620,091
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 491
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-620-091-81

Query Match      100.0%; Score 38; DB 4; Length 16;
Best Local Similarity 57.1%; Pred. No. 31;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 1 GYWCW 7

RESULT 3
US-09-543-681A-7620
; Sequence 7620, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7620
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7620

Query Match      100.0%; Score 38; DB 4; Length 177;
Best Local Similarity 57.1%; Pred. No. 2.9e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
```

```
Db 43 GYWINW 49
||||:|
US-09-502-653-10
; Sequence 10, Application US/09502653
; Patent No. 6331426
; GENERAL INFORMATION:
; APPLICANT: Bj invad, Mads Eskelund
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Schlein, Martin
; APPLICANT: Bech, Lisbeth
; APPLICANT: stergaard, Peter Rahbek
; APPLICANT: Sj holm, Carsten
; TITLE OF INVENTION: NOVEL GALACTANASES
; FILE REFERENCE: 5481.200-US
; CURRENT APPLICATION NUMBER: US/09/502,653
; CURRENT FILING DATE: 2000-02-11
; EARLIER APPLICATION NUMBER: PA 1999 00184
; EARLIER FILING DATE: 1999-02-11
; EARLIER APPLICATION NUMBER: PA 1999 00799
; EARLIER FILING DATE: 1999-06-07
; EARLIER APPLICATION NUMBER: 60/125,885
; EARLIER FILING DATE: 1999-03-24
; EARLIER APPLICATION NUMBER: 60/138,445
; EARLIER FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Bacillus circulans
US-09-502-653-10

Query Match      100.0%; Score 38; DB 3; Length 252;
Best Local Similarity 57.1%; Pred. No. 4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 210 GYWPW 216
||||:|

RESULT 5
US-09-610-906-12
; Sequence 12, Application US/09610906
; Patent No. 6566066
; GENERAL INFORMATION:
; APPLICANT: Walker, Michael G.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Klinger, Tod M.
; TITLE OF INVENTION: AQUAPORIN-8 VARIANT
; FILE REFERENCE: PC-0012 CIP
; CURRENT APPLICATION NUMBER: US/09/610,906
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 09/226,994
; PRIOR FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. 6566066: g2346968
; PUBLICATION INFORMATION:
US-09-610-906-12

Query Match      100.0%; Score 38; DB 4; Length 263;
Best Local Similarity 57.1%; Pred. No. 4.2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GYWXXXW 7
Db      224 GYWDFFW 230

RESULT 6
US-09-248-796A-15791
; Sequence 15791, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15791
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15791

Query Match      100.0%; Score 38; DB 4; Length 274;
Best Local Similarity 57.1%; Pred. No. 4.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 GYWXXXW 7
Db      84 GYWPITW 90

RESULT 7
US-09-145-828A-11
; Sequence 11, Application US/09145828A
; Patent No. 6403349
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Thurmond, Jennifer
; APPLICANT: Kirchner, Stephen J.
; APPLICANT: Parker-Barnes, Jennifer M.
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
; FILE REFERENCE: 6407.US.01
; CURRENT APPLICATION NUMBER: US/09/145,828A
; CURRENT FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-145-828A-11

Query Match      100.0%; Score 38; DB 4; Length 278;
Best Local Similarity 57.1%; Pred. No. 4.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 GYWXXXW 7
Db      108 GYWIFLW 114

RESULT 8
US-09-903-456-18
; Sequence 18, Application US/09903456
```

```
; Patent No. 6677145
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-903-456-18

Query Match      100.0%; Score 38; DB 4; Length 278;
Best Local Similarity 57.1%; Pred. No. 4.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 GYWXXXW 7
Db      108 GYWIFLW 114

RESULT 9
US-09-252-991A-26841
; Sequence 26841, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26841
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26841

Query Match      100.0%; Score 38; DB 4; Length 339;
Best Local Similarity 57.1%; Pred. No. 5.3e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 GYWXXXW 7
Db      261 GYWGSIW 267

RESULT 10
US-09-252-991A-32031
; Sequence 32031, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
```

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32031
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32031

Query Match 100.0%; Score 38; DB 4; Length 342;
Best Local Similarity 57.1%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXW 7
DB 315 GYWRGPW 321

RESULT 11
US-08-415-751-6
; Sequence 6, Application US/08415751
; Patent No. 5643772
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
; TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA
; TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
; TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
; TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS AND KIT
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PHILLIPS, MOORE, LEMPJO & FINLEY
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: California
; COUNTRY: United States of America
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
; COMPUTER: PC
; OPERATING SYSTEM: DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/415,751
; FILING DATE: 03-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/071,880
; FILING DATE: June 1, 1993
; APPLICATION NUMBER: 07/891,301
; FILING DATE: May 29, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hana Dolezalova
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.19-2 (HHD)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-1677
; TELEFAX: (415) 324-1678
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Cryptosporidium parvum
; FEATURE:
; NAME/KEY: Positions coded by nonsense codons are
; NAME/KEY: identified as Xaa.
US-08-415-751-6

Query Match 100.0%; Score 38; DB 1; Length 362;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXW 7
DB 216 GYWLTW 222

RESULT 12
US-09-248-796A-15188
; Sequence 15188, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15188
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15188

Query Match 100.0%; Score 38; DB 4; Length 367;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXW 7
DB 276 GYWLVDW 282

RESULT 13
US-09-107-532A-6868
; Sequence 6868, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:


```
/ APPLICATION NUMBER: 60/085,598
/ FILING DATE: 14 May 1998
/ APPLICATION NUMBER: 60/051571
/ FILING DATE: July 2, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ariniello, Pamela Deneke
/ REGISTRATION NUMBER: 40,489
/ REFERENCE/DOCKET NUMBER: GTC-012
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781)893-5007
/ TELEFAX: (781)893-8277
/ INFORMATION FOR SEQ ID NO: 6868:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 478 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: YES
/ ORIGINAL SOURCE:
/ ORGANISM: Enterococcus faecium
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (B) LOCATION 1...478
/ SEQUENCE DESCRIPTION: SEQ ID NO: 6868:
US-09-107-532A-6868

Query Match 100.0%; Score 38; DB 4; Length 478;
Best Local Similarity 57.1%; Pred. No. 7.3e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
DB 98 GYWLSAW 104

RESULT 14
US-09-107-532A-6945
/ Sequence 6945, Application US/09107532A
/ Patent No: 6583275
/ GENERAL INFORMATION:
/ APPLICANT: Lynn A Doucette-Stamm and David Bush
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
/ ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
/ NUMBER OF SEQUENCES: 7310
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: GENOME THERAPEUTICS CORPORATION
/ STREET: 100 Beaver Street
/ CITY: Waltham
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02354
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD-ROM ISO9660
/ COMPUTER: PC
/ OPERATING SYSTEM: <Unknown>
/ SOFTWARE: ASCII
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/107,532A
/ FILING DATE: 30-Jun-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/085,598
/ FILING DATE: 14 May 1998
/ APPLICATION NUMBER: 60/051571
/ FILING DATE: July 2, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ariniello, Pamela Deneke
/ REGISTRATION NUMBER: 40,489
/ REFERENCE/DOCKET NUMBER: GTC-012
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781)893-5007
/ TELEFAX: (781)893-8277
/ INFORMATION FOR SEQ ID NO: 6945:
/ SEQUENCE CHARACTERISTICS:
```

```
/ LENGTH: 492 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: YES
/ ORIGINAL SOURCE:
/ ORGANISM: Enterococcus faecium
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (B) LOCATION 1...492
/ SEQUENCE DESCRIPTION: SEQ ID NO: 6945:
US-09-107-532A-6945
```

```
Query Match 100.0%; Score 38; DB 4; Length 492;
Best Local Similarity 57.1%; Pred. No. 7.5e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GYWXW 7
DB 112 GYWLTCW 118
```

```
RESULT 15
US-09-252-991A-23328
/ Sequence 23328, Application US/09252991A
/ Patent No: 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 23328
/ LENGTH: 499
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23328
```

```
Query Match 100.0%; Score 38; DB 4; Length 499;
Best Local Similarity 57.1%; Pred. No. 7.6e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GYWXW 7
DB 111 GYWLAW 117
```

```
Search completed: January 3, 2005, 16:33:49
Job time : 21.16 secs
```

THIS PAGE BLANK (USPIC)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:32:45 ; Search time 70.56 Seconds
(without alignments)

35.687 Million cell updates/sec

Title: US-10-046-922-67

Perfect score: 38

Sequence: 1 GYVXXXW 7

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	7	13	US-10-046-922-67
2	38	100.0	8	13	US-10-046-922-67
3	38	100.0	9	13	US-10-046-922-67
4	38	100.0	10	13	US-10-046-922-34
5	38	100.0	10	13	US-10-046-922-35
6	38	100.0	10	13	US-10-046-922-73
7	38	100.0	13	14	US-10-125-869A-73
8	38	100.0	13	15	US-10-462-282-297
9	38	100.0	25	14	US-10-280-066-476
10	38	100.0	29	9	US-09-864-761-35127
11	38	100.0	35	9	US-09-864-761-40409
12	38	100.0	47	17	US-10-425-115-287762
13	38	100.0	57	16	US-10-437-963-170197

14	38	100.0	58	17	US-10-425-115-308836	Sequence 308836,
15	38	100.0	61	17	US-10-425-115-324628	Sequence 324628,
16	38	100.0	62	17	US-10-425-115-362818	Sequence 362818,
17	38	100.0	72	16	US-10-437-963-176036	Sequence 176036,
18	38	100.0	82	15	US-10-424-599-240891	Sequence 240891,
19	38	100.0	104	17	US-10-425-115-273234	Sequence 273234,
20	38	100.0	110	15	US-10-424-599-252989	Sequence 252989,
21	38	100.0	113	14	US-10-369-493-19264	Sequence 19264, A
22	38	100.0	116	10	US-09-910-483-1	Sequence 1, Appli
23	38	100.0	116	10	US-09-910-483-5	Sequence 5, Appli
24	38	100.0	116	10	US-09-910-483-9	Sequence 9, Appli
25	38	100.0	116	10	US-09-910-483-13	Sequence 13, Appli
26	38	100.0	116	10	US-09-910-483-17	Sequence 17, Appli
27	38	100.0	116	10	US-09-910-483-21	Sequence 21, Appli
28	38	100.0	116	10	US-09-910-483-25	Sequence 25, Appli
29	38	100.0	116	10	US-09-910-483-29	Sequence 29, Appli
30	38	100.0	116	10	US-09-910-483-33	Sequence 33, Appli
31	38	100.0	116	10	US-09-910-483-37	Sequence 37, Appli
32	38	100.0	116	10	US-09-910-483-41	Sequence 41, Appli
33	38	100.0	116	10	US-09-910-483-43	Sequence 43, Appli
34	38	100.0	122	14	US-10-447-331-6	Sequence 6, Appli
35	38	100.0	123	16	US-10-437-963-173556	Sequence 173556,
36	38	100.0	126	17	US-10-425-115-260243	Sequence 260243,
37	38	100.0	129	17	US-10-425-115-357518	Sequence 357518,
38	38	100.0	138	14	US-10-160-232-86	Sequence 86, Appli
39	38	100.0	138	14	US-10-160-232-90	Sequence 90, Appli
40	38	100.0	153	15	US-10-282-122A-68730	Sequence 68730, A
41	38	100.0	161	16	US-10-787-701-54778	Sequence 54778, A
42	38	100.0	166	15	US-10-425-114-64112	Sequence 64112, A
43	38	100.0	185	17	US-10-808-807-12	Sequence 12, Appli
44	38	100.0	199	15	US-10-424-599-267810	Sequence 267810,
45	38	100.0	227	15	US-10-282-122A-64263	Sequence 64263, A

ALIGNMENTS

RESULT 1

US-10-046-922-67
; Sequence 67, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 67
; LENGTH: 7
; TYPE: PRT
; ORGANISM: peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)..(6)
; OTHER INFORMATION: X at position 4-6 is any amino acid
US-10-046-922-67

Query Match 100.0%; Score 38; DB 13; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+06; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Oy 1 GYVXXXW 7

Db 1 GYVXXXW 7

RESULT 2

US-10-046-922-68
; Sequence 68, Application US/10046922

```
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 68
; LENGTH: 8
; TYPE: PRT
; ORGANISM: peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)..(6)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (8)..(8)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-68

Query Match
Best Local Similarity 100.0%; Score 38; DB 13; Length 8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXX 7
DB 1 GYXXXXX 7

RESULT 3
US-10-046-922-36
; Sequence 36, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 9
; TYPE: PRT
; ORGANISM: peptide
US-10-046-922-36

Query Match
Best Local Similarity 100.0%; Score 38; DB 13; Length 9;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXX 7
DB 2 GYXXDTW 8

RESULT 4
US-10-046-922-34
; Sequence 34, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73
; LENGTH: 10
; TYPE: PRT
; ORGANISM: peptide library
; FEATURE:

; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 10
; TYPE: PRT
; ORGANISM: isolated peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(1)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (10)..(10)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-34

Query Match
Best Local Similarity 100.0%; Score 38; DB 13; Length 10;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXX 7
DB 2 GYWLTIW 8

RESULT 5
US-10-046-922-35
; Sequence 35, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35
; LENGTH: 10
; TYPE: PRT
; ORGANISM: isolated peptide
US-10-046-922-35

Query Match
Best Local Similarity 100.0%; Score 38; DB 13; Length 10;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXX 7
DB 2 GYWLTIW 8

RESULT 6
US-10-046-922-73
; Sequence 73, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73
; LENGTH: 10
; TYPE: PRT
; ORGANISM: peptide library
; FEATURE:
```

; NAME/KEY: SITE
; LOCATION: (5)..(7)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (9)..(9)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-73

Query Match 100.0%; Score 38; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXW 7
Db 2 GYXXXXXW 8

RESULT 7

US-10-125-869A-73
; Sequence 73, Application US/10125869A
; Publication No. US20030199671A1
; GENERAL INFORMATION:
; APPLICANT: Rondon, Isaac Jesus
; APPLICANT: Wu, Qi-Long
; APPLICANT: Ley, Arthur C.
; APPLICANT: Stochl, Mark
; APPLICANT: Ranschoff, Thomas C.
; APPLICANT: Potter, M. Daniel (deceased)
; TITLE OF INVENTION: BINDING MOLECULES FOR Fc-REGION
; FILE REFERENCE: 3421.1006-001
; CURRENT APPLICATION NUMBER: US/10/125,869A
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/284,534
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fc region binding polypeptide
US-10-125-869A-73

Query Match 100.0%; Score 38; DB 14; Length 13;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXW 7
Db 1 GYWCNVW 7

RESULT 8

US-10-462-262-297
; Sequence 297, Application US/10462262
; Publication No. US20040009534A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Aaron K.
; APPLICANT: Dawson, Bruce M.
; TITLE OF INVENTION: PROTEIN ANALYSIS
; FILE REFERENCE: 10280-052001
; CURRENT APPLICATION NUMBER: US/10/462,262
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/388,642
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 297
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: immunoglobulin binding polypeptide
US-10-462-262-297

Query Match 100.0%; Score 38; DB 15; Length 13;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXW 7
Db 1 GYWCNVW 7

RESULT 9

US-10-280-066-476
; Sequence 476, Application US/10280066
; Publication No. US20030180718A1
; GENERAL INFORMATION:
; APPLICANT: Pillutla, Renuka C.
; APPLICANT: Briessette, Renee
; APPLICANT: Spruyt, Michael
; APPLICANT: Dedova, Olga
; APPLICANT: Blume, Arthur J.
; APPLICANT: Prendergast, John
; APPLICANT: Goldstein, Neil I.
; TITLE OF INVENTION: TARGET SPECIFIC SCREENING AND ITS USE FOR IDENTIFYING TARGET BINDING
; FILE REFERENCE: 2598-4009US1
; CURRENT APPLICATION NUMBER: US/10/280,066
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 60/345,471
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 537
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 476
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Eschericia coli
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: Tietl-20C-3-D116
US-10-280-066-476

Query Match 100.0%; Score 38; DB 14; Length 25;
Best Local Similarity 57.1%; Pred. No. 2.6e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXW 7
Db 6 GYWGELW 12

RESULT 10

US-09-864-761-35127
; Sequence 35127, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmics-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35127
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009503.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91
; OTHER INFORMATION: EST_HUMAN HIT: AA343827.1, EVALUATE 1.80e+00
US-09-864-761-35127

Query Match 100.0%; Score 38; DB 9; Length 29;
Best Local Similarity 57.1%; Pred. No. 2.9e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 17 GYWLHW 23

RESULT 11
US-09-864-761-40409
; Sequence 40409, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40409
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009503.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EST_HUMAN HIT: AI248042.1, EVALUATE 9.00e-03
US-09-864-761-40409

Query Match 100.0%; Score 38; DB 9; Length 35;
Best Local Similarity 57.1%; Pred. No. 3.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 25 GYWLHW 31

RESULT 12
US-10-425-115-287762
; Sequence 287762, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 287762
; LENGTH: 47

; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_25533C.1.pep
US-10-425-115-287762

Query Match 100.0%; Score 38; DB 17; Length 47;
Best Local Similarity 57.1%; Pred. No. 4.3e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 40 GYWTIF 46
|||::|

RESULT 13
US-10-437-963-170197
; Sequence 170197, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 170197
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68546C.1.pep
US-10-437-963-170197

Query Match 100.0%; Score 38; DB 16; Length 57;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 19 GYWLPM 25
|||::|

RESULT 14
US-10-425-115-308836
; Sequence 308836, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 308836
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_44726C.1.pep
US-10-425-115-308836

Query Match 100.0%; Score 38; DB 17; Length 58;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 49 GYWGASW 55
|||::|

RESULT 15
US-10-425-115-324628
; Sequence 324628, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 324628
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_59132C.1.pep
US-10-425-115-324628

Query Match 100.0%; Score 38; DB 17; Length 61;
Best Local Similarity 57.1%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 41 GYWKQW 47
|||::|

Search completed: January 3, 2005, 16:54:14
Job time : 70.56 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:20:13 ; Search time 15.68 Seconds
(without alignments)
42.954 Million cell updates/sec

Title: US-10-046-922-67
Perfect score: 38
Sequence: 1 GYXXXXW 7

Scoring table: _BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Préd. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	71	2 S22905	lysis protein S -
2	38	100.0	72	2 T03190	hypothetical prote
3	38	100.0	83	2 E69903	hypothetical prote
4	38	100.0	108	2 S12193	hypothetical prote
5	38	100.0	142	2 C34903	Ig heavy chain pre
6	38	100.0	187	2 G83047	hypothetical prote
7	38	100.0	218	2 S76385	hypothetical prote
8	38	100.0	218	2 D87264	hypothetical prote
9	38	100.0	227	2 S73905	CDPdiacylglycerol-
10	38	100.0	250	2 A69843	hypothetical prote
11	38	100.0	254	2 F82733	arginine-TRNA-prot
12	38	100.0	257	2 E75325	probable mccc prot
13	38	100.0	261	2 JCS806	aquaporin 8 - mous
14	38	100.0	263	2 JCS622	aquaporin 8 - rat
15	38	100.0	271	2 F81188	phosphatidate cytl
16	38	100.0	271	2 JC4832	phosphatidate cytl
17	38	100.0	273	2 E95268	probable ABC trans
18	38	100.0	279	2 AB2307	hypothetical prote
19	38	100.0	282	2 H95869	probable sugar ABC
20	38	100.0	286	2 E88690	protein F4H10.7 [
21	38	100.0	289	2 G72215	oligopeptide ABC t
22	38	100.0	309	2 C83886	hypothetical prote
23	38	100.0	344	2 C82611	hypothetical prote
24	38	100.0	345	2 T37139	hypothetical prote
25	38	100.0	360	2 AE2047	hypothetical prote
26	38	100.0	421	2 D82500	hypothetical prote
27	38	100.0	441	2 C95307	probable transport
28	38	100.0	447	2 H97146	siderophore/Surfac
29	38	100.0	448	2 AB0301	conserved hypothet

30	38	100.0	466	2 T35164	probable secreted
31	38	100.0	469	2 D70048	ABC transporter (a
32	38	100.0	472	2 E83497	probable amino aci
33	38	100.0	475	2 T48745	arginine/ornithine
34	38	100.0	482	2 JH0110	arginine/ornithine
35	38	100.0	490	2 C86879	arginine/ornithine
36	38	100.0	497	2 G86878	arginine/ornithine
37	38	100.0	508	2 C95282	probable ABC trans
38	38	100.0	517	2 A13201	hypothetical prote
39	38	100.0	519	2 S77572	oligopeptide trans
40	38	100.0	519	2 E83268	probable carboxydr
41	38	100.0	534	2 T15414	hypothetical prote
42	38	100.0	535	2 B95952	probable dipeptide
43	38	100.0	536	2 G95389	probable ABC trans
44	38	100.0	541	2 AC2392	hypothetical prote
45	38	100.0	563	2 AH2975	hypothetical prote

ALIGNMENTS

RESULT 1

S22905
lysis protein S - phase 21
C:Species: phase 21
C>Date: 09-Jun-1994 #sequence_revision 10-Nov-1995 #text_change 08-Oct-1999
C:Accession: S22905
R:Bonovich, M.T.; Young, R.
J. Bacteriol. 173, 2897-2905, 1991
A>Title: Dual start motif in two lambdoid S genes unrelated to lambda S.
A:Reference number: S22905; MUID:91210180; PMID:2019562
A:Accession: S22905
A:Molecule type: DNA
A:Residues: 1-71 <BON>
A:Cross-references: EMBL:M65239; NID:G215466; PIDN:AAA32349.1; PID:G215467
C:Genetics:

A:Gene: S
C:Keywords: alternative initiators; cell wall lysis; transmembrane protein
F:1-71/Product: lysis protein inhibitor S107 #status predicted <MAT1>
F:4-71/Product: lysis protein S105 #status predicted <MAT2>
F:7-28/Domain: transmembrane #status predicted <TM1>
F:36-58/Domain: transmembrane #status predicted <TM2>

Query Match 100.0%; Score 38; DB 2; Length 71;
Best Local Similarity 57.1%; Pred. No. 52;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GYXXXXW 7
Db	21	GYWFQW 27

RESULT 2

T03190
hypothetical protein 72B - rice mitochondrion
C:Species: mitochondrion Oryza sativa (rice)
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C:Accession: T03190
R:Itadani, H.; Wakasugi, T.; Sugita, M.; Sugiura, M.; Nakazono, M.; Hirai, A.
Plant Cell Physiol. 35, 1239-1244, 1994
A>Title: Nucleotide sequence of a 28-kbp portion of rice mitochondrial DNA: the existence
A:Reference number: Z14841; MUID:95211382; PMID:7545979
A:Accession: T03190
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-72 <TA>

A:Cross-references: UNIPROT:Q35302; EMBL:D32052; NID:G769704; PIDN:BAA06811.1; PID:G76970
A:Experimental source: cultivar Nipponbare
C:Genetics:

Query Match 100.0%; Score 38; DB 2; Length 72;

```
Best Local Similarity 57.1%; Pred. No. 52;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 34 GYWSHW 40

RESULT 3
hypothetical protein yodI - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: E69903
R:Kunat, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
A.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Bieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Setowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Toato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580, MUID:98044033; PMID:9384377
A:Accession: E69903
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-83 <KUN>
A:Cross-references: UNIPROT:Q34654; GB:299114; GB:AL009126; NID:G2634230; PIDN:CAB13852.
A:Experimental source: strain 168
A:Genetics:
A:Gene: yodI

Query Match 100.0%; Score 38; DB 2; Length 83;
Best Local Similarity 57.1%; Pred. No. 60;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 53 GYWGWW 59

RESULT 4
S12193
hypothetical protein 4 - Thiobacillus ferrooxidans plasmid pTF1
C:Species: Thiobacillus ferrooxidans
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S12193
R:Drolet, M.; Zanga, P.; Lau, P.C.K.
Mol. Microbiol. 4, 1381-1391, 1990
A:Title: The mobilization and origin of transfer regions of a Thiobacillus ferrooxidans
A:Reference number: S12188; MUID:91125140; PMID:2280689
A:Accession: S12193
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-108 <DRO>
A:Cross-references: UNIPROT:P20088; EMBL:X52699; NID:G48158; PIDN:CAA36930.1; PID:G48164
C:Genetics:
A:Genome: plasmid pTF1

Query Match 100.0%; Score 38; DB 2; Length 108;
Best Local Similarity 57.1%; Pred. No. 76;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 89 GYWRSSW 95
```

RESULT 5

```
C34903
IG heavy chain precursor V region (5-27) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 16-Aug-1996
C:Accession: C34903
R:Bedzyk, W.D.; Herron, J.N.; Edmundson, A.B.; Voss Jr., E.W.
J. Biol. Chem. 265, 133-138, 1990
A:Title: Active site structure and antigen binding properties of idiotypically cross-react
A:Reference number: A34903; MUID:90094387; PMID:2104617
A:Accession: C34903
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-142 <BED>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:34-119/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 38; DB 2; Length 142;
Best Local Similarity 57.1%; Pred. No. 97;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 126 GYWFYW 132

RESULT 6
GB3047
hypothetical protein PA4793 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: GB3047
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: GB3047
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-187 <STO>
A:Cross-references: UNIPROT:Q9HV15; GB:AE004892; GB:AE004091; NID:99951049; PIDN:AAG0817;
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA4793

Query Match 100.0%; Score 38; DB 2; Length 187;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
Db 109 GYGGYW 115

RESULT 7
S76385
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S76385
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76385
```

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <KAN>
A:Cross-references: UNIPROT:Q55705; EMBL:D64000; GB:AB001339; NID:g1001484; PIDN:BAAL023
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: probable alkaline phosphatase yngC

Query Match 100.0%; Score 38; DB 2; Length 218;
Best Local Similarity 57.1%; Pred. No. 1.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
DB 74 GYWGGRW 80
||||:|

RESULT 8
D87264
hypothetical protein CC0125 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: D87264
R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
N.; J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: D87264
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <STO>
A:Cross-references: UNIPROT:Q9ABU5; GB:AE005673; NID:g13421234; PIDN:AAK22112.1; GSPDB:C
C:Genetics:
A:Gene: CC0125

Query Match 100.0%; Score 38; DB 2; Length 218;
Best Local Similarity 57.1%; Pred. No. 1.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
DB 139 GYWRPAW 145
||||:|

RESULT 9
S73905
CDPdiacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase (EC 2.7.8.5) pgSA - Myc
N:Alternate names: hypothetical protein A65_orf227
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342
C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S73905
R.Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A:Reference number: S73327; MUID:97105885; PMID:8948633
A:Accession: S73905
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-227 <HIM>
A:Cross-references: UNIPROT:P75520; EMBL:AE000057; GB:U00089; NID:g1674279; PIDN:AA89622
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetics:
A:Gene: pgSA
A:Genetic code: SGC3
C:Superfamily: CDPdiacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase
C:Keywords: transferase

Query Match 100.0%; Score 38; DB 2; Length 227;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
DB 84 GYWARCW 90
||||:|

RESULT 10
A69843
hypothetical protein yjba - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: A69843
R.Kunat, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallert
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoisappel, S.; Hosono, S.; Hullo, M.F.;
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel,
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Bieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seror,
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: A69843
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-250 <KUN>
A:Cross-references: UNIPROT:O31597; GB:Z99110; GB:AL009126; NID:g2633472; PIDN:CAB12998.1
A:Experimental source: strain 168
C:Genetics:
A:Gene: yjba
C:Superfamily: Bacillus subtilis hypothetical protein yjba

Query Match 100.0%; Score 38; DB 2; Length 250;
Best Local Similarity 57.1%; Pred. No. 1.6e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
DB 185 GYWTGW 191
||||:|

RESULT 11
F82733
arginine-tRNA-protein transferase XP1018 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: F82733
R.anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: F82733
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-254 <SIM>
A:Cross-references: UNIPROT:O9PEL0; GB:AE003939; GB:AE003849; NID:g9105949; PIDN:AAF8382
R.Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; AJ
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.
as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohne
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laigre
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A:Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1018

Query Match 100.0%; Score 38; DB 2; Length 254;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
DB 22 GYWPDRW 28

RESULT 12

E75325
Probable mcfF protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: E75325
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: E75325
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-257 <WHI>
A:Cross-references: UNIPROT:Q9RSX1; GB:AE002038; GB:AE000513; NID:96459790; PIDN:AAF1155
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2000
A:Map position: 1

Query Match 100.0%; Score 38; DB 2; Length 257;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
DB 171 GYWLNRW 177

RESULT 13

JCS806
aquaporin 8 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 27-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C:Accession: JCS806
R:Ma, T.; Yang, B.; Verkman, A.S.
Biochem. Biophys. Res. Commun. 240, 324-328, 1997
A:Title: Cloning of a novel water and urea-permeable aquaporin from mouse expressed str
A:Reference number: JCS806; MUID:98049830; PMID:9388476
A:Accession: JCS806
A:Molecule type: mRNA
A:Residues: 1-261 <MAA>
A:Cross-references: UNIPROT:P56404; DBJ:AF018952; NID:92353796; PIDN:AA568847.1; PID:92
C:Comment: This protein functions as a mercurial-sensitive water channel.
C:Superfamily: lens fiber membrane major intrinsic protein
C:Keywords: glycoprotein
F:92-94/Region: NPA motif
F:210-212/Region: NPA motif
F:85,139/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 38; DB 2; Length 261;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7

DB 222 GYWDPHW 228

RESULT 14

JCS622
aquaporin 8 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 09-Oct-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C:Accession: JCS622
R:Ishibashi, K.; Kuwahara, M.; Kageyama, Y.; Tohseaka, A.; Marumo, F.; Sasaki, S.
Biochem. Biophys. Res. Commun. 237, 714-718, 1997
A:Title: Cloning and functional expression of a second new aquaporin abundantly expressed
A:Reference number: JCS622; MUID:97445104; PMID:9299432
A:Contents: Testis
A:Accession: JCS622
A:Molecule type: mRNA
A:Residues: 1-263 <ISH>
A:Cross-references: UNIPROT:P56405; DBJ:AB005547; NID:92346967; PIDN:BAA21918.1; PID:92
C:Comment: This protein is a water channel protein which plays a role in the regulation
C:Superfamily: lens fiber membrane major intrinsic protein
C:Keywords: glycoprotein
F:39-59/Domain: transmembrane #status predicted <TM1>
F:65-84/Domain: transmembrane #status predicted <TM2>
F:94-96/Region: NPA motif
F:109-130/Domain: transmembrane #status predicted <TM3>
F:159-179/Domain: transmembrane #status predicted <TM4>
F:182-204/Domain: transmembrane #status predicted <TM5>
F:212-214/Region: NPA motif
F:231-250/Domain: transmembrane #status predicted <TM6>
F:141/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 38; DB 2; Length 263;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
DB 224 GYWDPHW 230

RESULT 15

F83188
phosphatidate cytidyltransferase PA3651 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F83188
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: F83188
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-271 <STO>
A:Cross-references: UNIPROT:Q59640; GB:AE004785; GB:AE004091; NID:99949809; PIDN:AAG0703;
A:Experimental source: strain PA01
C:Genetics:
A:Gene: cdsA; PA3651
C:Superfamily: phosphatidate cytidyltransferase

Query Match 100.0%; Score 38; DB 2; Length 271;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
DB 102 GYWGGRW 108

Search completed: January 3, 2005, 16:34:52

Job time : 16.68 secs



THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using ew model

Run on: January 3, 2005, 16:12:00 ; Search time 85.12 Seconds
(without alignments)
47.317 Million cell updates/sec

Title: US-10-046-922-67
Perfect score: 38
Sequence: 1 GYVXXW 7

Scoring table: .BLOSUM62DX.*
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	30	2 Q7UDB7	Q7udb7 shigella fl
2	38	100.0	49	2 Q9EVP1	Q9evp1 escherichia
3	38	100.0	71	1 VLXS BPP21	P27360 bacteriophage
4	38	100.0	71	2 Q7C2J0	Q7c2j0 shigella fl
5	38	100.0	71	2 Q9FCW3	Q9fcw3 escherichia
6	38	100.0	71	2 Q7UDP0	Q7udp0 shigella fl
7	38	100.0	71	2 Q83ML2	Q83ml2 shigella fl
8	38	100.0	71	2 Q83S57	Q83s57 shigella fl
9	38	100.0	72	2 Q35302	Q35302 oryza sativ
10	38	100.0	83	1 Y0BI BACSU	Q34654 bacillus su
11	38	100.0	89	2 Q95S05	Q95s05 drosophila
12	38	100.0	102	2 Q6IIL4	Q6iil4 drosophila
13	38	100.0	108	1 YML2 THIFE	P20088 thicobacillu
14	38	100.0	120	2 Q728A6	Q728a6 desulfovibr
15	38	100.0	120	2 AAS97170	Aas97170 desulfovi
16	38	100.0	122	2 Q72D02	Q72d02 desulfovibr
17	38	100.0	122	2 AAS95609	Aas95609 desulfovi
18	38	100.0	124	2 Q7U7V6	Q7u7v6 synchococc
19	38	100.0	128	2 Q8MK57	Q8mk57 bos taurus
20	38	100.0	130	2 Q7U395	Q7u395 prochloroco
21	38	100.0	130	2 Q7VBG3	Q7vbg3 prochloroco
22	38	100.0	135	2 Q7TUV7	Q7tuv7 prochloroco
23	38	100.0	160	2 Q6NPF17	Q6nfp17 corynebacte
24	38	100.0	160	2 CAES0615	Caes0615 corynebac
25	38	100.0	187	2 Q9HV15	Q9hv15 pseudomonas
26	38	100.0	189	2 Q8Q0B8	Q8q0b8 pseudomonas
27	38	100.0	204	2 Q7W0P5	Q7w0p5 bordetella
28	38	100.0	204	2 Q7W3F9	Q7w3f9 bordetella
29	38	100.0	204	2 Q7WES9	Q7wes9 bordetella
30	38	100.0	212	2 Q7WL18	Q7wl18 bordetella
31	38	100.0	213	2 Q9KY37	Q9ky37 streptomyce

32 38 100.0 214 2 Q8A8U4
33 38 100.0 218 1 Y232 SINY3
34 38 100.0 218 2 Q7VZM7
35 38 100.0 218 2 Q9ABU5
36 38 100.0 224 2 Q8S486
37 38 100.0 227 1 PGSA_MYCPN
38 38 100.0 228 2 Q72KR5
39 38 100.0 228 2 AAS80801
40 38 100.0 235 2 Q7W7N0
41 38 100.0 236 2 Q9KGW9
42 38 100.0 243 2 Q8CX61
43 38 100.0 250 2 Q31597
44 38 100.0 253 2 Q32816
45 38 100.0 254 1 ATE_XYLFA

Q8A8U4 bacteroides
Q55705 synchocyst
Q7vzm7 bordetella
Q9ABU5 caulobacter
Q8S486 zea mays (m
P75520 mycoplasma
Q72kr5 thermus t
Aas80801 thermus t
Q7w7n0 bordetella
Q9kgw9 human immun
Q8cx61 arabidopsis
Q31597 bacillus su
Q32816 lactococcus
Q9psl0 xylella fas

ALIGNMENTS

RESULT 1
Q7UDB7 PRELIMINARY; PRT; 30 AA.
AC Q7UDB7;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical bacteriophage protein.
GN Name=ybcR; OrderedLocustNames=S0714;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner P.R.;
RT "Complete genome sequence and comparative genomics of Shigella
flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786 (2003).
DR EMBL; AE016980; AAP16193.1; -.
DR InterPro; IPR007054; Lysis_S.
DR Pfam; PF04971; Lysis_S; 1.
KW Hypothetical protein.
SQ SEQUENCE 30 AA; 3404 MW; 7EA4C66BE5C1486E CRC64;

Query Match 100.0%; Score 38; DB 2; Length 30;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYVXXW 7
|||:::
Db 21 GYVFLQW 27

RESULT 2
Q9EVP1 PRELIMINARY; PRT; 49 AA.
ID Q9EVP1;
AC Q9EVP1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE S protein (Fragment).
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H.I.8.;

```

RX MEDLINE=20407286; PubMed=10948097;
RA Unkneir A.; Schmidt H.;
RT "Structural analysis of phage-borne stx genes and their flanking
RT sequences in shiga toxin-producing Escherichia coli and Shigella
RT dysenteriae type 1 strains.";
RL Infect. Immun. 68:4856-4864(2000).
DR EMBL; AJ271139; CAC05573.1; -.
DR InterPro; IPR007054; Lysis_S.
DR PFam; PF04971; Lysis_S; 1.
FT NON TER 49
SQ SEQUENCE 49 AA; 5227 MW; 0B6914DD9AE25E00 CRC64;

Query Match 100.0%; Score 38; DB 2; Length 49;
Best Local Similarity 57.1%; Pred. No. 2.6e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
DB 21 GYWFLQW 27

RESULT 3
VLXS_BPP21 STANDARD; PRT; 71 AA.
AC P27360;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lysis protein S.
GN Name=S;
OS Bacteriophage P21 (Bacteriophage 21).
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10711;
RN [1]
RP MEDLINE=91210180; PubMed=2019562;
RX Bonovich M.T., Young R.;
RT "Dual motif in two lambdaoid S genes unrelated to lambda S.";
RL J. Bacteriol. 173:2897-2905(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M65239; AAA32349.1; -.
DR InterPro; IPR007054; Lysis_S.
DR PFam; PF04971; Lysis_S; 1.
KW Phage lysis protein.
SQ SEQUENCE 71 AA; 7893 MW; 8690A6F25234A3E2 CRC64;

Query Match 100.0%; Score 38; DB 1; Length 71;
Best Local Similarity 57.1%; Pred. No. 3.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
DB 21 GYWFLQW 27

RESULT 4
Q7C2J0 PRELIMINARY; PRT; 71 AA.
AC Q7C2J0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative S protein.
GN OrderedLocusNames=S0731;

```

```

OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AF016980; AAP16205.1; -.
DR InterPro; IPR007054; Lysis_S.
DR PFam; PF04971; Lysis_S; 1.
SQ SEQUENCE 71 AA; 7914 MW; 10CE1C485234AE99 CRC64;

Query Match 100.0%; Score 38; DB 2; Length 71;
Best Local Similarity 57.1%; Pred. No. 3.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
DB 21 GYWFLQW 27

RESULT 5
Q9FCW3 PRELIMINARY; PRT; 71 AA.
AC Q9FCW3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE S protein.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T4/97;
RX MEDLINE=20407286; PubMed=10948097;
RA Unkneir A., Schmidt H.;
RT "Structural analysis of phage-borne stx genes and their flanking
RT sequences in shiga toxin-producing Escherichia coli and Shigella
RT dysenteriae type 1 strains.";
RL Infect. Immun. 68:4856-4864(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=T4/97;
RA Unkneir A., Karch H., Schmidt H.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=T4/97;
RA Schmidt H., Scheef J., Morabito S., Caprioli A., Wieler L., Karch H.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ270998; CAC05565.1; -.
DR InterPro; IPR007054; Lysis_S.
DR PFam; PF04971; Lysis_S; 1.
SQ SEQUENCE 71 AA; 7923 MW; 9B4D68F25220B7E2 CRC64;

Query Match 100.0%; Score 38; DB 2; Length 71;
Best Local Similarity 57.1%; Pred. No. 3.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
DB 21 GYWFLQW 27

```



```
RESULT 6
Q7UDP0
ID Q7UDP0 PRELIMINARY; PRT; 71 AA.
AC Q7UDP0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 26, Last annotation update)
DE Lysis protein S.
GN OrderedLocusNames=S0231;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; A016978; AAP15759.1; -.
DR InterPro; IPR007054; Lysis_S.
DR Pfam; PF04971; Lysis_S; 1.
DR [1]
SQ SEQUENCE 71 AA; 7881 MW; AB82BAF25234BC15 CRC64;

Query Match 100.0%; Score 38; DB 2; Length 71;
Best Local Similarity 57.1%; Pred. No. 3.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXX 7
DB 21 GYWFLQW 27

RESULT 7
Q83ML2
ID Q83ML2 PRELIMINARY; PRT; 71 AA.
AC Q83ML2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE LYSIS PROTEIN S.
GN OrderedLocusNames=SF2038;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL; A015099; AAN42324.1; -.
DR InterPro; IPR007054; Lysis_S.
DR Pfam; PF04971; Lysis_S; 1.
DR [1]
SQ SEQUENCE 71 AA; 7881 MW; AB82BAF25234BC15 CRC64;

Query Match 100.0%; Score 38; DB 2; Length 71;
Best Local Similarity 57.1%; Pred. No. 3.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXX 7
DB 21 GYWFLQW 27

RESULT 7
Q83ML2
ID Q83ML2 PRELIMINARY; PRT; 71 AA.
AC Q83ML2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE LYSIS PROTEIN S.
GN OrderedLocusNames=SF2038;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL; A015220; AAN43581.1; -.
DR InterPro; IPR007054; Lysis_S.
DR Pfam; PF04971; Lysis_S; 1.
DR [1]
SQ SEQUENCE 71 AA; 7865 MW; AB90A8F25234A3F5 CRC64;

Query Match 100.0%; Score 38; DB 2; Length 71;
Best Local Similarity 57.1%; Pred. No. 3.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GYXXXXX 7
DB 21 GYWFLQW 27

RESULT 8
Q83S57
ID Q83S57 PRELIMINARY; PRT; 71 AA.
AC Q83S57;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Lysis protein S.
GN OrderedLocusNames=SF0689;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL; A015099; AAN42324.1; -.
DR InterPro; IPR007054; Lysis_S.
DR Pfam; PF04971; Lysis_S; 1.
DR [1]
SQ SEQUENCE 71 AA; 7914 MW; 10CE1C485234AE99 CRC64;

Query Match 100.0%; Score 38; DB 2; Length 71;
Best Local Similarity 57.1%; Pred. No. 3.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXX 7
DB 21 GYWFLQW 27

RESULT 9
Q35302
ID Q35302 PRELIMINARY; PRT; 72 AA.
AC Q35302;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ORF72B.
OS Oryza sativa (japonica cultivar-group).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Baev A.A., Dzhumagaliev E.B., Lyubomirskaya N.V., Mizrokhi L.Y.,
RA Il'in Y.V.;
RT "Structure of long and short copies of the mobile dispersed gene MDG3
RT of Drosophila melanogaster."
RL Dokl. Akad. Nauk SSSR 282:1483-1486(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95308541; PubMed=7788722;
RA Nakazono M., Itadani H., Wakasugi T., Tsutsumi N., Sugitara M.,
RA Hirai A.;
RT "The rps3-rpl16-nad3-rps12 gene cluster in rice mitochondrial DNA is
```

```

RT transcribed from alternative promoters.";
RL Curr. Genet. 27:184-189(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95211382; PubMed=7545979;
RA Itadani H., Wakasugi T., Sugita M., Sugiura M., Nakazono M., Hirai A.;
RT "Nucleotide sequence of a 28-kbp portion of rice mitochondrial DNA:
RT the existence of many sequences that correspond to parts of
RT mitochondrial genes in intergenic regions.";
RL Plant Cell Physiol. 35:1239-1244(1994).
DR EMBL; D32052; BAA06811.1; -.
DR PIR; T03190; T03190.
DR Gramene; Q35302; -.
DR GO; GO:005739; C:mitochondrion; IEA.
DR Mitochondrion.
SQ SEQUENCE 72 AA; 8155 MW; D42DE53BED28432E CRC64;
Query Match 100.0%; Score 38; DB 2; Length 72;
Best Local Similarity 57.1%; Pred. No. 3.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYWXXXW 7
DB 34 GYWSHW 40
[1]:::
RESULT 10
YODI BACSU STANDARD; PRT; 83 AA.
AC Q34654;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical protein yodi.
GN Names=yodi; Synonyms=yola; OrderedLocusNames=BSU19610;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX Wambutt R., Wedler H., Lapidus A., Sorokin A., Ehrlich S.D.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
RT the odhA and aspC loci cloned in a yeast artificial chromosome.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Ghim S.-Y., Jeong Y.-M., Choi S.-K., Park S.-H.;
RT "Sequence analysis of the 30 kb region (192') of the Bacillus subtilis
RT chromosome containing the cse cluster.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Brouillet S., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Choi S.-K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N.,
RA Ghim S.-Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood S.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullio M.F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,

```

```

RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieser M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeder R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weitzneger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL Nature 390:249-256(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF015775; AAB72056.1; -.
DR EMBL; AF006665; AAB81166.1; -.
DR EMBL; Z99114; CAB13852.1; -.
DR PIR; E69903; E69903.
DR Subtilist; BG13537; yodi.
DR InterPro; IPR008991; Transl SH3 like.
KW Complete proteome; Hypothetical protein; Transmembrane.
FT TRANSMEM 58 Potential.
SQ SEQUENCE 83 AA; 9194 MW; 99F58EA2F0F36A43 CRC64;
Query Match 100.0%; Score 38; DB 1; Length 83;
Best Local Similarity 57.1%; Pred. No. 4.2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYWXXXW 7
DB 53 GYGGYW 59
[1]:::
RESULT 11
Q9SS05 PRELIMINARY; PRT; 89 AA.
ID Q9SS05;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE HL03793p.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Ceiniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY061016; AAL28564.1; -.
DR FlyBase; FBgn0047239; BGDNA:HL03793.
SQ SEQUENCE 89 AA; 10063 MW; 36CB86917DF80B9D CRC64;
Query Match 100.0%; Score 38; DB 2; Length 89;
Best Local Similarity 57.1%; Pred. No. 4.5e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYWXXXW 7

```

```

Db          70 GYWCQNW 76
||||:|
RESULT 12
Q6IIL4      PRELIMINARY;      PRT;      102 AA.
AC Q6IIL4;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE HDC17696.
GN ORFNames=HDC17696;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14709175;
RA Hild M., Beckmann B., Haas S., Koch B., Solovyev V., Busold C.,
RA Fellenberg K., Boutros M., Vingron M., Sauer F., Honeisel J., Paro R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome.";
RL Genome Biol. 5:R3-R3(2003).
CC -|- MISCELLANEOUS (The sequence shown here is derived from an
CC EMBL/GenBank/DDJB third party annotation (TFA) entry.
DR EMBL; BK003052; DAA03252.1; -.
SQ SEQUENCE 102 AA; 11777 MW; CD4044EC9325CF46 CRC64;

Query Match          100.0%; Score 38; DB 2; Length 102;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
||||:|
DB 73 GYWIQEW 79

RESULT 13
YML2_THIFE STANDARD;      PRT;      108 AA.
AC P20088;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 12.3 kDa protein in mobL 3' region (ORF 4).
OS Thiobacillus ferrooxidans.
OG Plasmid pTF1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Acidithiobacillales;
OC Acidithiobacillaceae; Acidithiobacillus.
OX NCBI_TaxID=920;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33020;
RX MEDLINE=91125140; PubMed=2280689;
RA Drollet M., Zanga P., Lau P.C.K.;
RT "The mobilization and origin of transfer regions of a Thiobacillus
RT ferrooxidans plasmid; relatedness to plasmids RSF1010 and pSCI01.";
RL Mol. Microbiol. 4:1381-1391(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52699; CAA36930.1; -.
DR PIR; S12193; S12193.
KW Hypothetical protein; Plasmid.

```

```

SQ SEQUENCE 108 AA; 12335 MW; A8E67717C109A57E CRC64;

Query Match          100.0%; Score 38; DB 1; Length 109;
Best Local Similarity 57.1%; Pred. No. 5.3e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
||||:|
DB 89 GYWRSSW 95

RESULT 14
Q728A6      PRELIMINARY;      PRT;      120 AA.
AC Q728A6;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Lipoprotein, putative.
GN OrderedlocusNames=DVU2698;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Daviden T.M., Zafar N., Zhou L., Radune D.,
RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AE017318; AAS97170.1; -.
DR TIGR; DVU2698; -.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 120 AA; 13418 MW; CD3581657D76E183 CRC64;

Query Match          100.0%; Score 38; DB 2; Length 120;
Best Local Similarity 57.1%; Pred. No. 5.9e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXW 7
||||:|
DB 54 GYWDWRW 60

RESULT 15
AAS97170    PRELIMINARY;      PRT;      120 AA.
AC AAS97170;
DT 26-APR-2004 (TReMBLrel. 27, Created)
DT 26-APR-2004 (TReMBLrel. 27, Last sequence update)
DT 11-MAY-2004 (TReMBLrel. 27, Last annotation update)
DE Lipoprotein, putative.
GN DVU2698.
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15077118;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,

```

RA Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,
RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AE017318; AAS97170.1; -.
DR TIGR; DVU2698; -.
KW Lipoprotein.
SQ SEQUENCE 120 AA; 13418 MW; CD3581657D76E183 CRC64;

Query Match 100.0%; Score 38; DB 2; Length 120;
Best Local Similarity 57.1%; Pred. No. 5.9e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXW 7
Db 54 GYWIDRW 60

Search completed: January 3, 2005, 16:32:33
Job time : 88.12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 17:00:47 ; Search time 146 Seconds
(without alignments)
17.199 Million cell updates/sec

Title: US-10-046-922-67

Perfect score: 38

Sequence: 1 GYXXXXW 7

Scoring table: -BLOSUM62DX¹
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 116873

Minimum DB seq length: 0

Maximum DB-seq-length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	7	3 AAY76794	Aay76794 Somatosta
2	38	100.0	7	5 ABP53964	Abp53964 VEGFR-3 b
3	38	100.0	7	5 ABP53418	Abp53418 Backbone
4	32	84.2	6	2 AAR93713	Aar93713 Cyclo[-Iy
5	32	84.2	6	5 AAU83934	Aau83934 Tyrosine
6	32	84.2	7	2 AAR08140	Aar08140 Neurokini
7	32	84.2	7	3 AAY76792	Aay76792 Somatosta
8	32	84.2	7	5 ABP53416	Abp53416 Backbone
9	32	84.2	7	8 ADJ25834	Adj25834 Tyrosine
10	28	73.7	6	2 AAR48807	Aar48807 Peptide w
11	28	73.7	6	2 AAR46290	Aar46290 PTP-DI/D2
12	28	73.7	6	2 AAW43021	Aaw43021 Protein t
13	28	73.7	6	3 AAY81787	Aay81787 Protein t
14	28	73.7	7	2 AAR62543	Aar62543 Fibronect
15	28	73.7	7	2 AAY14751	Aay14751 Tyrosine
16	28	73.7	7	2 AAY14752	Aay14752 Tyrosine
17	28	73.7	7	2 AAY33139	Aay33139 Rabbit ca
18	27	71.1	6	2 AAW28937	Aaw28937 Opioide pe
19	27	71.1	6	2 AAR93795	Aar93795 New pepi
20	27	71.1	6	2 AAY23046	Aay23046 Opioide pe
21	27	71.1	6	4 AAG78945	Aag78945 Antimicro
22	27	71.1	6	8 ABG75386	Abg75386 C melo se
23	27	71.1	7	2 AAW40924	Aaw40924 Leader se
24	27	71.1	7	2 AAY16913	Aay16913 Heat shoc
25	27	71.1	7	2 AAY16925	Aay16925 Heat shoc

26	27	71.1	7	3 AAY76793	Aay76793 Somatosta
27	27	71.1	7	3 AAU78246	Aau78246 Clonorch
28	27	71.1	7	3 AAU77087	Aau77087 Clonorch
29	27	71.1	7	4 AAU72055	Aau72055 Melanoma
30	27	71.1	7	4 AAU72067	Aau72067 Melanoma
31	27	71.1	7	5 AAU80617	Aau80617 Javelin p
32	27	71.1	7	5 AAU80605	Aau80605 Javelin p
33	27	71.1	7	5 AAU81571	Aau81571 Enterokin
34	27	71.1	7	5 ABP53417	Abp53417 Backbone
35	27	71.1	7	5 AAU70373	Aau70373 Mouse hea
36	27	71.1	7	6 ABU79853	Abu79853 Enterokin
37	27	71.1	7	6 AAO26659	Aao26659 chit36 de
38	27	71.1	7	7 ADG72569	Adg72569 Heat shoc
39	27	71.1	7	7 ADG72581	Adg72581 Heat shoc
40	27	71.1	7	7 ADG72777	Adg72777 Heat shoc
41	27	71.1	7	7 ADG72789	Adg72789 Heat shoc
42	27	71.1	7	7 ADL17115	Adl17115 Phage-dis
43	27	71.1	7	7 ADL17125	Adl17125 Phage-dis
44	26	68.4	5	2 AAR76079	Aar76079 MAB 55.1
45	26	68.4	5	3 AAY32257	Aay32257 Light cha

ALIGNMENTS

RESULT 1

AA76794
ID AAY76794 standard; peptide; 7 AA.

XX AC AAY76794;
DT 20-APR-2000 (first entry)
XX
DE Somatostatin analogue peptide 3181.

KW Somatostatin analogue; therapy; cyclic peptide; autoimmune disease;
KW endocrine disorder; cancer; diabetic-associated complication; diagnosis;
KW gastrointestinal disorder; inflammatory disease; pancreatitis;
KW atherosclerosis; restenosis; post-surgical pain; VIP secretion inhibitor;
KW hormone-secreting tumour; hormone-dependent tumour; diarrhoea;
KW vasoactive intestinal peptide; non-insulin dependent diabetes mellitus.
OS Synthetic.

PH Key Location/Qualifiers
FT Misc-difference 3 /note= "D-form residue"
FT Modified-site 7 /note= "Trp-NH2"
FT
XX WO9965508-A1.
XX 23-DEC-1999.
XX 15-JUN-1999; 99WO-IL000329.
XX 19-JUN-1998; 98US-00100360.
XX 02-DEC-1998; 98US-00203389.
(PEPT-) PEPTOR LTD.
XX Hornik V, Afargan MM, Gellerman G;
DR WPI; 2000-136888/12.
XX
XX Cyclic somatostatin analogs for inhibiting growth hormone secretion
XX from anterior pituitary and as antiproliferative agents for the treatment
XX of tumors.
XX Example 11; Page 61; 82pp; English.
XX This sequence represents a somatostatin analogue of the invention. The
XX invention relates to a backbone cyclised somatostatin analogue that has

CC one building unit containing a nitrogen atom of the peptide backbone
 CC connected to a bridging group comprising an amide, thioether, thioester
 CC or disulphide. At least one building unit is connected via a bridging
 CC group to form a cyclic structure with a moiety selected from a second
 CC building unit, side chain of or N-terminal amino acid residue. A
 CC composition containing the analogue may be used for preventing disorders
 CC such as cancers, autoimmune diseases, endocrine disorders, diabetic-
 CC associated complications, gastrointestinal disorders, inflammatory
 CC diseases, pancreatitis, atherosclerosis, restenosis and post-surgical
 CC pain. It may also be used for diagnosing cancer. The backbone cyclic
 CC analogue is used for imaging the existence of metastases. Somatostatin
 CC analogues can be used for the treatment of patients with hormone-secreting
 CC and hormone-dependent tumours. They reduce diarrhoea through the
 CC inhibition of vasoactive intestinal peptide (VIP) secretion and by direct
 CC effect on intestinal secretion. Somatostatin analogues selective to type
 CC 2 and 5 receptors may be used for treatment of non-insulin dependent
 CC diabetes mellitus. They are useful for the prevention of atherosclerosis
 CC and restenosis. The analogues are metabolically stable, selective in
 CC their in-vivo activities and safe
 XX Sequence 7 AA;

Query Match 100.0%; Score 38; DB 3; Length 7;
 Best Local Similarity 57.1%; Pred. No. 1.7e+06;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
 |||::|
 Db 1 GYWKVCW 7

RESULT 2
 ABP53964
 ID ABP53964 standard; peptide; 7 AA.
 XX
 AC ABP53964;
 XX
 DT 09-JAN-2003 (first entry)
 XX
 DE VEGFR-3 binding peptide SEQ ID NO:67.
 KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
 KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
 KW cytostatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
 KW vulnary; cell surface receptor; cancer; neovascularisation;
 KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
 KW diabetes; PDGF; platelet derived growth factor.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 4. .6 /note= "X is any amino acid"
 FT
 XX WO200257299-A2.
 XX
 XX 25-JUL-2002.
 XX
 XX 16-JAN-2002; 2002WO-IB000099.
 XX
 XX 17-JAN-2001; 2001US-0262476P.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX (LICN) LICENTIA LTD.
 XX
 XX Alitalo K, Koivunen E, Kubo H;
 XX
 XX WPI; 2002-691521/74.
 DR
 XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
 PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
 PT such as cancer and diseases of neovascularization.

XX
 PS Claim 21; Page 81; 149pp; English.
 XX
 CC The present invention describes an isolated peptide (I) that binds to and
 CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnerary activities, and can be used in gene therapy.
 CC Compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
 CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 XX Sequence 7 AA;

Query Match 100.0%; Score 38; DB 5; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
 |||::|
 Db 1 GYWXXXW 7

RESULT 3
 ABP53418
 ID ABP53418 standard; peptide; 7 AA.
 XX
 AC ABP53418;
 XX
 DT 19-NOV-2002 (first entry)
 XX
 DE Backbone cyclised somatostatin analogue PTR 3181.
 XX
 KW Backbone cyclised somatostatin analogue; somatostatin; SRIF; analgesic;
 KW somatotropin release inhibiting factor; somatostatin receptor subtype;
 KW sythesis; antiarteriosclerotic; immunosuppressive; cytostatic; cancer;
 KW antidiabetic; antiinflammatory; somatostatin receptor ligand;
 KW atherosclerosis; autoimmune disease; diabetic-associated complication;
 KW endocrine disorder; inflammation; gastrointestinal disorder; restenosis;
 KW pancreatitis; post-surgical pain.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminally modified with Fmoc
 (fluorenylmethoxycarbonyl)"
 FT Misc-difference 3 /note= "D form residue"
 FT Modified-site 7 /note= "amidated"
 FT
 XX US2002052315-A1.
 XX
 XX 02-MAY-2002.
 XX
 XX 13-DEC-2000; 2000US-00734583.
 XX
 XX 19-JUN-1998; 98US-00100360.
 XX 02-DEC-1998; 98US-00203389.
 XX 15-JUN-1999; 99WO-IL000329.
 XX (HORN/) HORNIK V.
 XX (AFAR/) AFARGAN M M.
 XX (GELL/) GELLERMAN G.
 XX
 XX Hornik V, Afargan MM, Gellerman G;

XX WPI; 2002-681319/73.
 XX New backbone cyclized somatostatin analogs are e.g. useful in the
 PT treatment of atherosclerosis, autoimmune diseases and cancers.
 XX
 XX Example 12; Page 21; 30pp; English.
 XX
 CC The present invention describes backbone cyclised somatostatin analogues
 CC (I) that incorporates at least one building unit containing one nitrogen
 CC atom of the peptide backbone connected to a bridging group (comprising an
 CC amide, thioether, thioester or disulfide) where at least one building
 CC unit is connected via the bridging group to form a cyclic structure with
 CC a moiety selected from the group consisting of a second building unit,
 CC the side chain of an amino acid residue of the sequence or the N-terminal
 CC amino acid residue. (I) has antiarteriosclerotic, immunosuppressive,
 CC cytostatic, antidiabetic, antiinflammatory and analgesic activities, and
 CC can be used as a somatostatin receptor ligand. (I) are useful in the
 CC treatment of atherosclerosis, autoimmune diseases, cancers, diabetic-
 CC associated complications, endocrine disorders, inflammation,
 CC gastrointestinal disorders, pancreatitis, post-surgical pain, and
 CC restenosis. (I) can also be used in the diagnosis of cancer, by imaging
 CC the existence of metastases, it being labeled with a detectable probe.
 CC The present sequence represents a backbone cyclised somatostatin analogue
 CC from the present invention
 XX
 XX Sequence 7 AA;
 SQ
 Query Match 100.0%; Score 38; DB 5; Length 7;
 Best Local Similarity 57.1%; Pred. No. 1.7e+06;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GYVXXW 7
 DB |||::|
 1 GYVXCV 7
 RESULT 4
 AAR93713
 ID AAR93713 standard; peptide; 6 AA.
 AC AAR93713;
 XX
 XX 10-MAY-1996 (first entry)
 DT
 DE Cyclo[-Tyr-trp-Leu-Arg-Gly-Trp-].
 XX
 KW neurokinin A antagonist; tachykinin; respiratory disease; asthma;
 KW analgesic; cyclic.
 XX
 OS Synthetic.
 XX
 FT Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "not an N-terminal amino acid, but condensed with
 FT Trp(6) to form a cyclic peptide"
 FT Misc-difference 2
 FT /note= "D-form residue"
 FT Modified-site 6
 FT /note= "not a C-terminal amino acid, but condensed with
 FT Tyr(1) to form a cyclic peptide"
 FT
 PN W0521187-A1.
 XX
 PD 10-AUG-1995.
 XX
 XX 10-JAN-1995; 95WO-US000296.
 PF
 PR 03-FEB-1994; 94US-00191571.
 XX
 XX (RICH) MERRELL DOW PHARM INC.
 PA
 XX Owen TJ, Kudlacz EM, Buck SH, Harbeson SL;
 PI

XX WPI; 1995-336695/43.
 XX
 PT New cyclic peptide derivs. - are neurokinin A and tachykinin antagonists
 PT useful e.g. for treating asthma or as analgesics.
 XX
 XX Claim 8; Page 69; 82pp; English.
 XX
 CC The patent describes novel cyclic hexapeptide and octapeptide compounds
 CC which are antagonists of neurokinin A and which are useful medically as
 CC analgesics and for treating respiratory diseases such as asthma. The
 CC present sequence represents a specifically preferred example of the new
 CC peptides
 XX
 XX Sequence 6 AA;
 SQ
 Query Match 84.2%; Score 32; DB 2; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.7e+06;
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YWXXW 7
 DB |||::|
 1 YWLRGW 6
 RESULT 5
 AAU83934
 ID AAU83934 standard; peptide; 6 AA.
 XX
 AC AAU83934;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE Tyrosine recombinase inhibitory peptide #37.
 XX
 KW Tyrosine recombinase; antibacterial; cytostatic; cell growth modulator;
 KW site-specific DNA recombinase; type I DNA topoisomerase; tumour;
 KW tyrosine recombinase; cancer; Holliday junction.
 XX
 OS Synthetic.
 XX
 PN W0200198540-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 21-JUN-2001; 2001WO-US020046.
 XX
 PR 22-JUN-2000; 2000US-00602087.
 XX
 PA (UUSA-) UNIV SAN DIEGO STATE FOUND.
 XX
 PI Segall A, Pinilla C;
 XX
 DR WPI; 2002-114591/15.
 XX
 PT Identifying cell growth modulators for inhibiting cancer cell growth in
 PT humans, involves assessing and comparing activity of site-specific
 PT recombinase or type I DNA topoisomerase in presence/absence of test
 PT substance.
 XX
 PS Claim 39; Page 90; 115pp; English.
 XX
 CC The method relates to a method of identifying a modulator of cell growth,
 CC comprising assessing and comparing the activities of site-specific DNA
 CC recombinase (I) or type I DNA topoisomerase (II) in presence and absence
 CC of a test substance. A difference in activity of (I) and (II) assessed in
 CC the presence and absence of the test substance indicates that the test
 CC substance modulates cell growth. The identified cell growth modulator,
 CC preferably an inhibitor of (I) or (II), is useful for inhibiting cell
 CC growth in a subject, preferably a human. The inhibitor inhibits (I) which
 CC is preferably tyrosine recombinase or type I DNA topoisomerase in humans
 CC having or suspected of having tumour or cancer, where the method further
 CC involves administering an effective of antitumour or anticancer agent or

CC treatment; or who are, or are suspected of being infected by a bacterium,
CC in which case the inhibitor inhibits Holliday junction intermediate
CC resolution activity of tyrosine recombinase. The method further involves
CC administering an effective amount of antibiotic or antibacterium
CC treatment. AAU8398-AAU83991 represents tyrosine recombinase inhibitory
CC peptides of the invention
XX
XX Sequence 6 AA;

Query Match 84.2%; Score 32; DB 5; Length 6;
Best Local Similarity 50.0%; Pred. NO. 1.7e+06;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 YWXXW 7
DB 1 YWCYWW 6

RESULT 6
ID AAR08140 standard; protein; 7 AA.
XX

AC AAR08140;
XX
DT 27-FEB-1991 (first entry)
XX
DE Neurokinin A peptide antagonist.
XX
XX NK-2-receptor; bronchoconstriction; spasms.
XX
OS Synthetic.

FH Key Location/Qualifiers
FT Modified-site 3 /label= D-Trp
FT Modified-site 5 /label= D-Trp
FT Modified-site 6 /label= D-Trp
FT Modified-site 7 /label= D-Trp
FT Modified-site 7 /label= D-Trp
XX
XX EP401177-A.

XX 05-DEC-1990.
XX
XX 25-MAY-1990; 90EP-00830234.
XX
XX 29-MAY-1989; 89IT-00009432.
XX
XX (MENA) MENARINI A IND FARM.
XX
XX Rovero P, Pestellini V, Maggi CA, Patacchini R, Santiccioli P;
PI Giuliani S, Meli A;
XX
XX WPI; 1990-363658/49.
XX
XX Synthetic peptide antagonists of neurokinin A - solid phase synthesis,
PT active against bronchoconstricting or spasms.
XX
XX Claim 1; Page 7; 9pp; English.
XX
XX Peptide antagonists of the NK-2-receptor of neurokinin A are useful in
CC the treatment of Bronchoconstrictions or spasms of the intestines or
CC urinary bladder
XX
XX Sequence 7 AA;

Query Match 84.2%; Score 32; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. NO. 1.7e+06;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 YWXXW 7

DB ||::|
2 YWXXW 7
RESULT 7
AAAY76792
ID AAY76792 standard; peptide; 7 AA.
XX
AC AAY76792;
XX
DT 20-APR-2000 (first entry)
XX
DE Somatostatin analogue peptide 3177.
XX

XX Somatostatin analogue; therapy; cyclic peptide; autoimmune disease;
KW endocrine disorder; cancer; diabetic-associated complication; diagnosis;
KW gastrointestinal disorder; inflammatory disease; pancreatitis;
KW atherosclerosis; restenosis; post-surgical pain; VIP secretion inhibitor;
KW hormone-secreting tumour; hormone-dependent tumour; diarrhoea;
KW vasoactive intestinal peptide; non-insulin dependent diabetes mellitus.
XX
OS Synthetic.

XX Key Location/Qualifiers
FH Misc-difference 3 /note= "D-form residue"
FT Modified-site 7 /note= "Trp-NH2"
FT
XX
XX W09965508-A1.
XX
XX 23-DEC-1999.
XX

XX 15-JUN-1999; 99WO-IL000329.
XX
XX 19-JUN-1998; 98US-00100360.
XX
XX 02-DEC-1998; 98US-00203389.
XX
XX (PEPT-) PEPTO LTD.

XX
XX Hornik V, Afargan MM, Gellerman G;

XX WPI; 2000-136888/12.

XX Cyclized somatostatin analogs for inhibiting growth hormone secretion
PT from anterior pituitary and as antiproliferative agents for the treatment
PT of tumors.

XX Example 11; Page 61; 82pp; English.

XX This sequence represents a somatostatin analogue of the invention. The
CC invention relates to a backbone cyclised somatostatin analogue that has
CC one building unit containing a nitrogen atom of the peptide backbone
CC connected to a bridging group comprising an amide, thioether, thioester
CC or disulphide. At least one building unit is connected via a bridging
CC group to form a cyclic structure with a moiety selected from a second
CC building unit, side chain of or N-terminal amino acid residue. A
CC composition containing the analogue may be used for preventing disorders
CC such as cancers, autoimmune diseases, endocrine disorders, diabetic-
CC associated complications, gastrointestinal disorders, inflammatory
CC diseases, pancreatitis, atherosclerosis, restenosis and post-surgical
CC pain. It may also be used for diagnosing cancer. The backbone cyclic
CC analogue is used for imaging the existence of metastases. Somatostatin
CC analogues can be used for the treatment of patients with hormone-secreting
CC and hormone-dependent tumours. They reduce diarrhoea through the
CC inhibition of vasoactive intestinal peptide (VIP) secretion and by direct
CC effect on intestinal secretion. Somatostatin analogues selective to type
CC 2 and 5 receptors may be used for treatment of non-insulin dependent
CC diabetes mellitus. They are useful for the prevention of atherosclerosis
CC and restenosis. The analogues are metabolically stable, selective in
XX their in-vivo activities and safe
XX Sequence 7 AA;

Query Match 84.2%; Score 32; DB 3; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.7e+06;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7
DB 2 YWKVCW 7

RESULT 8
ID ABP53416 standard; peptide; 7 AA.
XX AC ABP53416;
XX DT 19-NOV-2002 (first entry)
XX DE Backbone cyclised somatostatin analogue PTR 3177.
XX KW Backbone cyclised somatostatin analogue; somatostatin; SRIF; analgesic;
KW somatostatin release inhibiting factor; somatostatin receptor subtype;
KW synthesis; antiarteriosclerotic; immunosuppressive; cytostatic; cancer;
KW antidiabetic; antiinflammatory; somatostatin receptor ligand;
KW atherosclerosis; autoimmune disease; diabetic-associated complication;
KW endocrine disorder; inflammation; gastrointestinal disorder; restenosis;
KW pancreatitis; post-surgical pain.
XX OS Synthetic.
XX PH Key Location/Qualifiers
FT Misc-difference 3
FT /note= "D form residue"
FT Modified-site 7
FT /note= "amidated"

US2002052315-A1.
02-MAY-2002.
13-DEC-2000; 2000US-00734583.
19-JUN-1998; 98US-00100360.
02-DEC-1998; 98US-00203389.
15-JUN-1999; 99WO-IL000329.
(HORN/) HORNIK V.
(AFAR/) AFARGAN M M.
(GELL/) GELLERMAN G.
Hornik V, Afargan MM, Gellerman G;
WPI; 2002-681319/73.
New backbone cyclised somatostatin analogs are e.g. useful in the treatment of atherosclerosis, autoimmune diseases and cancers.
Example 12; Page 21; 30pp; English.

The present invention describes backbone cyclised somatostatin analogues (I) that incorporates at least one building unit containing one nitrogen atom of the peptide backbone connected to a bridging group (comprising an amide, thioether, thioester or disulfide) where at least one building unit is connected via the bridging group to form a cyclic structure with a moiety selected from the group consisting of a second building unit, the side chain of an amino acid residue of the sequence or the N-terminal amino acid residue. (I) has antiarteriosclerotic, immunosuppressive, cytostatic, antidiabetic, antiinflammatory and analgesic activities, and can be used as a somatostatin receptor ligand. (I) are useful in the treatment of atherosclerosis, autoimmune diseases, cancers, diabetic-associated complications, endocrine disorders, inflammation, gastrointestinal disorders, pancreatitis, post-surgical pain, and restenosis. (I) can also be used in the diagnosis of cancer, by imaging

CC the existence of metastases, it being labeled with a detectable probe.
CC The present sequence represents a backbone cyclised somatostatin analogue
CC from the present invention
XX Sequence 7 AA;
SQ Query Match 84.2%; Score 32; DB 5; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.7e+06;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7
DB 2 YWKVCW 7

RESULT 9
ID ADJ25834 standard; peptide; 7 AA.
XX AC ADJ25834;
XX DT 20-MAY-2004 (first entry)
XX DE Tyrosine tRNA synthetase binding peptide group 2 motif.
XX KW Tyrosine tRNA synthetase binding peptide group 2 motif.
KW ligand identification; peptide library;
KW complementary combinatorial library; tyrosine tRNA synthetase.
XX OS Synthetic.
XX PN US6617114-B1.
XX PD 09-SEP-2003.
XX PF 30-APR-1998; 98US-00069827.
XX PR 31-OCT-1996; 96US-00740671.
XX PR 31-OCT-1997; 97WO-US019638.
XX PR 31-MAR-1998; 98US-00050359.
XX PA (KARO-) KARO BIO AB.
XX PI Fowlkes DM, Kay BK, Frelinger JA, Hyde-Deruysscher RP;
XX WPI; 2004-068186/07.
XX PT Identification of ligand that can mediate biological activity of target protein, comprises screening first combinatorial library having first member ligands for binding to target protein to identify target-binding ligand(s).
XX PS Example 5; SEQ ID NO 94; 98pp; English.
XX CC The invention relates to a method of identifying a ligand that can mediate the biological activity of target protein via inhibition of the binding of target protein to a binding partner ligand comprising screening first combinatorial library having first member ligands for binding to target protein to identify target-binding ligand(s). The method is useful for identifying ligands that can mediate the biological activity of target proteins via inhibition of the binding of target protein to a binding partner ligand. The invention does not require that the natural binding partner be used as reagent. The need for the natural binding partner is obviated with the use of complementary combinatorial libraries. The present sequence is used in the exemplification of the present invention.

Sequence 7 AA;
SQ Query Match 84.2%; Score 32; DB 8; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.7e+06;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7

```

Db      ||::|
      1 YWHPDW 6

RESULT 10
AAR48807
ID AAR48807 standard; peptide; 6 AA.
XX
AC AAR48807;
XX
DT 25-MAR-2003 (revised)
XX
DT 12-AUG-1994 (first entry)
XX
DE Peptide which binds endothelin receptors.
XX
KW Endothelin; agonist; antagonist; peptide; receptor; cardiogenic shock;
KW hypertension; pulmonary hypertension; acute myocardial infarction;
KW uraemia; Crohn's disease; ulcerative colitis; sepsis; nephrotoxicity;
KW congestive heart failure; coronary spasm; cyclosporine.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1 /note= "D-form residue."
FT
FT
XX WO9403483-A1.
XX
PD 17-FEB-1994.
XX
PF 30-JUL-1993; 93WO-US007166.
XX
PR 30-JUL-1992; 92US-00923703.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Spellmeyer DC, Stauber GB, Simon RJ, Geysen HM;
XX WPI; 1994-065598/08.
XX
PT Endothelin receptor binding compounds - for treating hypertension,
PT cancer, psoriasis, atherosclerosis, hyperplasia, etc.
XX
PS Claim 8; Page 27; 40pp; English.
XX
CC vasculitis; pregnancy associated toxemia; transplantation; surgery.
CC Peptides which bind to endothelin receptors may act as agonists or
CC antagonists depending on whether they effect or block the activity of the
CC receptor. They may be used to modulate conditions such as cardiogenic
CC shock, hypertension, pulmonary hypertension, acute myocardial infarction,
CC uraemia, Crohn's disease and ulcerative colitis which are associated with
CC increased levels of endothelin. Increased endothelin levels are also
CC observed following orthotopic liver transplantation and major abdominal
CC surgical procedures. Endothelin may have a pathophysiologic role in
CC sepsis, congestive heart failure, coronary spasm, cyclosporine
CC nephrotoxicity, vasculitis and pregnancy associated toxemia. (See also
CC AAR4882-R48933). (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 6 AA;

Query Match 73.7%; Score 28; DB 2; Length 6;
Best Local Similarity 33.3%; Pred. No. 1.7e+06;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
Db :||:|
1 FWDIIW 6

RESULT 11
AAR46290
ID AAR46290 standard; protein; 6 AA.
XX
AC AAR46290;
XX
DT 25-MAR-2003 (revised)
XX
DT 26-AUG-1994 (first entry)
XX
DE PTP-D1/D2 primer #58 peptide.
XX
KW PTP-D1; PTP-D2; PCR; polymerase chain reaction; amplify; subfamily;
KW protein tyrosine phosphatase; PTP; catalytic domain; PTPase; antibody;
KW glycoprotein; cancer; diabetes; cellular; human;
KW phosphotyrosine metabolism.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 3 /note= "Undefined amino acid"
FT Misc-difference 5 /note= "Undefined amino acid"
FT
XX WO9403611-A2.
XX
PD 17-FEB-1994.
XX
PF 05-AUG-1993; 93WO-EP002086.
XX
PR 05-AUG-1992; 92US-00923740.
XX
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
PI Ullrich A, Moller N, Bach-Moller K;
XX WPI; 1994-065699/08.
XX
PT New protein tyrosine phosphatase(s), DNA encoding PTP-D and antibodies to
PT PTP-D - for identifying normal or mutant PTP-D genes, treating or
PT diagnosing e.g. cancer, diabetes or diseases associated with altered
PT tyrosine phosphate metabolism.
XX
PS Example 1; Page 43; 85pp; English.
XX
CC The sequences given in AAR46290-91 are peptides which were used to design
CC the primers given in AAQ57005-06 which were used in the amplification and
CC identification of the PTPases, PTP-D1 and PTP-D2. PTP-D1 and PTP-D2
CC represent members of a subfamily of protein tyrosine phosphatases (PTP)
CC which have slightly different structures from previously reported PTPs.
CC Members of the PTP-D subfamily are characterised by having one, two or
CC three identified amino acid changes in previously defined consensus
CC sequences in the catalytic phosphatase domains of known PTPases. PTP-D,
CC its antibodies or the DNA encoding PTP-D molecules may be used for
CC identifying normal or mutant PTP-D genes, or for measuring the amount or
CC activity of PTP-D protein or glycoprotein associated with a cell or
CC tissue. This may be utilised in methods for identifying susceptibility to
CC cancer, diabetes or other diseases associated with alterations in
CC cellular phospho- tyrosine metabolism. (Updated on 25-MAR-2003 to correct
CC PN field.)
XX
SQ Sequence 6 AA;

Query Match 73.7%; Score 28; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.7e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
Db :||:|
1 FWXMXW 6

RESULT 12
AAW43021
ID AAW43021 standard; peptide; 6 AA.
XX
AC AAW43021;

```

XX 06-JUL-1998 (first entry)
 XX Protein tyrosine phosphatase conserved peptide.
 DE
 XX
 KW Protein tyrosine phosphatase-neural and pancreatic; PTP-NP; mouse;
 KW cell growth; cell differentiation; cell survival; drug screening;
 KW tissue injury; wound healing; pancreas; neural cell; transgenic animal.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 5
 FT /label= Ile, Val
 XX
 PN WO9800552-A2.
 XX
 PD 08-JAN-1998.
 XX
 XX 27-JUN-1997; 97WO-US011519.
 PF
 XX
 PR 02-JUL-1996; 96US-0021040P.
 XX
 XX (HARD) HARVARD COLLEGE.
 PA
 XX Chiang M, Flanagan JG;
 PI
 XX WPI; 1998-086977/08.
 DR
 XX Protein tyrosine phosphatase-neural and pancreatic poly:peptide(s) and
 PT related nucleic acid - useful for modulating growth, differentiation and
 PT survival of cells and for drug screening.
 XX
 XX Example 1; Page 57; 99pp; English.
 PS
 XX
 CC This peptide corresponds to a conserved region of known protein tyrosine
 CC phosphatases (PTP). A sense primer (see AAV04937) based on this peptide
 CC was used in the PCR amplification of mouse neural tube cDNA. A 359 bp
 CC segment was isolated and used to screen a mouse brain cDNA library to
 CC identify several overlapping clones containing a 3.2 kb sequence (see
 CC AAV04936) encoding a novel PTP-NP (for neural and pancreatic) receptor
 CC (see AAW43023) that can be used for modulating growth, differentiation
 CC and survival of cells and for drug screening
 XX
 SQ Sequence 6 AA;
 Query Match 73.7%; Score 28; DB 2; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.7e+06;
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YWXXYW 7
 DB :|:|:
 1 FWRMXW 6
 RESULT 13
 AAY81787
 ID AAY81787 standard; peptide; 6 AA.
 XX
 AC AAY81787;
 XX
 XX 07-JUN-2000 (first entry)
 DT
 DE Protein tyrosine phosphatase motif #2.
 XX
 KW Human; protein tyrosine phosphatase; antibody; intracellular domain; LAR;
 KW CD45; PTP; diagnosis; insulin resistance related disease; syndrome X;
 KW non-insulin dependent diabetes mellitus; arteriosclerosis; therapy;
 KW heart disorder; signature motif.
 XX
 OS Unidentified.
 XX
 XX Key Location/Qualifiers
 FT

FT Misc-difference 3 /label= Arg, Glu, Leu
 FT Misc-difference 5 /label= Val, Ile, Cys
 FT
 XX WO200002922-A1.
 PN
 XX 20-JAN-2000.
 PD
 XX 06-JUL-1999; 99WO-JP003656.
 PP
 XX 10-JUL-1998; 98WO-JP003120.
 PR
 XX (FUSO) FUSO PHARM IND LTD.
 PA
 XX Yamamoto H, Tsujikawa K, Uchino Y;
 PI
 XX WPI; 2000-182215/16.
 DR
 XX Antibody for diagnosis and treatment of insulin resistance disorders and
 PT syndrome X recognises the intracellular domains of tyrosine kinases.
 PT
 XX Example 3; Page 32; 83pp; Japanese.
 PS
 XX This sequence represents a motif of a protein tyrosine phosphatases. The
 CC invention relates to an antibody of the invention that has sites
 CC specifically recognising the intracellular domain of two or more protein
 CC tyrosine phosphatases (PTPs). The antibody is useful for the detection
 CC and assay of PTP including novel phosphatases generated by cloning; and
 CC diagnosis, treatment and prevention of insulin resistance related
 CC diseases and non-insulin dependent diabetes mellitus, syndrome X and
 CC arteriosclerosis and heart disorders
 XX
 SQ Sequence 6 AA;
 Query Match 73.7%; Score 28; DB 3; Length 6;
 Best Local Similarity 66.7%; Pred. No. 1.7e+06;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YWXXYW 7
 DB :|:|:
 1 FWRMXW 6
 RESULT 14
 AAR62543
 ID AAR62543 standard; peptide; 7 AA.
 XX
 AC AAR62543;
 XX
 DT 25-MAR-2003 (revised)
 DT 16-DEC-1994 (first entry)
 XX
 DE Fibronectin gelatin binding domain inhibitory peptide.
 XX
 KW fibronectin; collagen-binding proteins; inhibit cell adhesion;
 KW regulate cell matrix interactions; tumorigenesis; metastasis;
 KW wound repair; homostasis; thrombospondin.
 XX
 OS Synthetic.
 XX
 XX WO9411395-A1.
 PN
 XX 26-MAY-1994.
 PD
 XX 09-NOV-1993; 93WO-US011104.
 PF
 XX 10-NOV-1992; 92US-00973235.
 PR
 XX (USSH) US SEC DEPT HEALTH.
 PA
 XX Roberts DD, Krutzh HC, Sipes JM, Guo N, Negre E;
 PI
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 17:14:38 ; Search time 37 Seconds
(without alignments)
12.547 Million cell updates/sec

Title: US-10-046-922-67
Perfect score: 38
Sequence: 1 GYVXXW 7

Scoring table: BL0SUM62DX⁷

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 58200

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/protdata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/protdata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/protdata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/protdata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/protdata/1/iaa/PCTUS_COMB.pep:*
- 6: /cgn2_6/protdata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	84.2	7	4	US-09-069-827A-94
2	29	76.3	7	1	US-08-443-640-16
3	28	73.7	6	2	US-08-446-345-12
4	28	73.7	6	3	US-08-951-260A-2
5	28	73.7	6	3	US-08-884-569A-11
6	28	73.7	6	3	US-08-884-569A-14
7	28	73.7	6	4	US-09-430-626A-2
8	28	73.7	6	4	US-09-361-096A-46
9	28	73.7	6	4	US-09-743-492A-5
10	28	73.7	6	4	US-10-243-687-2
11	28	73.7	7	1	US-07-973-235A-30
12	28	73.7	7	2	US-08-652-971-5
13	28	73.7	7	2	US-08-462-720-30
14	28	73.7	7	2	US-08-991-258A-5
15	28	73.7	7	2	US-08-769-399-5
16	28	73.7	7	3	US-08-991-953A-5
17	27	71.1	6	1	US-08-487-006-67
18	27	71.1	6	2	US-08-488-659A-67
19	27	71.1	7	4	US-09-563-222C-53
20	26	68.4	5	1	US-08-353-400-27
21	26	68.4	6	1	US-07-718-577-6
22	25	65.8	5	1	US-08-321-625-3
23	25	65.8	5	1	US-08-321-625-15
24	25	65.8	5	1	US-07-789-184-88
25	25	65.8	5	1	US-08-475-263-88
26	25	65.8	5	1	US-08-485-886-88
27	25	65.8	5	2	US-08-477-362-88

28	25	65.8	5	2	US-08-477-134-88	Sequence 88, Appl
29	25	65.8	5	3	US-08-473-489A-88	Sequence 88, Appl
30	25	65.8	5	3	US-08-485-695-88	Sequence 88, Appl
31	25	65.8	5	3	US-09-181-083-3	Sequence 3, Appl1
32	25	65.8	5	3	US-09-181-083-15	Sequence 15, Appl
33	25	65.8	5	3	US-08-018-760-88	Sequence 88, Appl
34	25	65.8	5	4	US-09-209-676-14	Sequence 14, Appl
35	25	65.8	5	4	US-09-535-852-2007	Sequence 2007, Ap
36	25	65.8	5	4	US-09-750-754-3	Sequence 3, Appl1
37	25	65.8	5	4	US-09-750-754-15	Sequence 15, Appl
38	25	65.8	5	5	PCT-US94-01321-48	Sequence 48, Appl
39	25	65.8	5	6	5185431-15	Patent No. 5185431
40	25	65.8	5	6	5217869-14	Patent No. 5217869
41	25	65.8	5	6	5217869-38	Patent No. 5217869
42	25	65.8	6	1	US-08-191-571-7	Sequence 7, Appl1
43	25	65.8	6	1	US-08-487-006-38	Sequence 38, Appl
44	25	65.8	6	1	US-08-439-817-191	Sequence 191, App
45	25	65.8	6	1	US-08-439-817-192	Sequence 192, App

ALIGNMENTS

RESULT 1
US-09-069-827A-94
; Sequence 94, Application US/09069827A
; Patent No. 6617114
; GENERAL INFORMATION:
; APPLICANT: FOWLKES, Dana M
; KAY, Brian K
; FRELINGER, Jeffrey A
; HYDE-DERUYSCHE, Robin P
; TITLE OF INVENTION: IDENTIFICATION OF DRUGS USING
; COMPLEMENTARY COMBINATORIAL LIBRARIES
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION NUMBER: US/09/069,827A
; FILING DATE: 30-Apr-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/050,359
; FILING DATE: 31-MAR-1998
; APPLICATION NUMBER: PCT/US97/19638
; FILING DATE: 31-OCT-1997
; APPLICATION NUMBER: US 08/740,671
; FILING DATE: 31-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Iver P
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: FOWLKES=4C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 94:

US-09-069-827A-94

Query Match 84.2%; Score 32; DB 4; Length 7;
Best Local Similarity 50.0%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
| | : : |
DB 1 YWPDW 6

RESULT 2

US-08-443-640-16
; Sequence 16, Application US/08443640
; Patent No. 5691140
; GENERAL INFORMATION:
; APPLICANT: EVANS, PAUL D.
; APPLICANT: NOREN, CHRISTOPHER J.
; TITLE OF INVENTION: BIDIRECTIONAL IN VITRO TRANSCRIPTION
; TITLE OF INVENTION: VECTORS UTILIZING A SINGLE RNA POLYMERASE FOR BOTH
; DIRECTIONS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEW ENGLAND BIOLABS, INC
; STREET: 32 TOZER ROAD
; CITY: BEVERLY
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 01915
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/443,640
; FILING DATE: 18-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.
; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: NEB-102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 927-5054
; TELEFAX: (508) 927-1705
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-443-640-16

Query Match 76.3%; Score 29; DB 1; Length 7;
Best Local Similarity 42.9%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXXW 7
| | : : |
DB 1 GWRFAW 7

RESULT 3

US-08-446-345-12
; Sequence 12, Application US/08446345
; Patent No. 5831009
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; TITLE OF INVENTION: NOVEL PROTEIN PHOSPHOTYROSINE
; PHOSPHATASES PTP-D1

NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: N.Y.
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,345
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/234,440
; FILING DATE: 28-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30742
; REFERENCE/DOCKET NUMBER: 7683-054
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-446-345-12

Query Match 73.7%; Score 28; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
| | : : |
DB 1 FWXXW 6

RESULT 4

US-08-951-260A-2
; Sequence 2, Application US/08951260A
; Patent No. 6004791
; GENERAL INFORMATION:
; APPLICANT: Aoki, Naohito
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PROTEIN TYROSINE PHOSPHATASE PTP20
; TITLE OF INVENTION: AND RELATED PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951,260A
; FILING DATE: October 16, 1997

```
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,860
; FILING DATE: No. 6004791ember 13, 1996
; APPLICATION NUMBER: PCT/1897/00946
; FILING DATE: June 17, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: "Xaa" in positions 3 and 5 stand
; OTHER INFORMATION: for an unspecified amino acid.
US-08-951-260A-2

Query Match 73.7%; Score 28; DB 3; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7
DB 1 FWXXW 6

RESULT 5
US-08-884-569A-11
; Sequence 11, Application US/08884569A
; Patent No. 6399326
; GENERAL INFORMATION:
; APPLICANT: CHIANG, MING-KO
; TITLE OF INVENTION: RECEPTOR TYROSINE PHOSPHATASE, AND USES RELATED THERETO
; FILE REFERENCE: HMV-020.01
; CURRENT APPLICATION NUMBER: US/08/884,569A
; CURRENT FILING DATE: 1997-06-27
; PRIOR APPLICATION NUMBER: 60/021,040
; PRIOR FILING DATE: 1996-07-02
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: MOD_RES
; LOCATION: (5)
; OTHER INFORMATION: Ile or Val
US-08-884-569A-11

Query Match 73.7%; Score 28; DB 3; Length 6;
Best Local Similarity 50.0%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7
DB 1 FWXXW 6

RESULT 6
US-08-884-569A-14
; Sequence 14, Application US/08884569A
; Patent No. 6399326
; GENERAL INFORMATION:
; APPLICANT: CHIANG, MING-KO
; TITLE OF INVENTION: RECEPTOR TYROSINE PHOSPHATASE, AND USES RELATED THERETO
; FILE REFERENCE: HMV-020.01
; CURRENT APPLICATION NUMBER: US/08/884,569A
; CURRENT FILING DATE: 1997-06-27
; PRIOR APPLICATION NUMBER: 60/021,040
; PRIOR FILING DATE: 1996-07-02
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: MOD_RES
; LOCATION: (3)
; OTHER INFORMATION: Arg or Gln
; NAME/KEY: MOD_RES
; LOCATION: (5)
; OTHER INFORMATION: Ile or Val
US-08-884-569A-14

Query Match 73.7%; Score 28; DB 3; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7
DB 1 FWXXW 6

RESULT 7
US-09-430-626A-2
; Sequence 2, Application US/09430626A
; Patent No. 6482605
; GENERAL INFORMATION:
; APPLICANT: Aoki, Naohito
; Ullrich, Axel
; TITLE OF INVENTION: PROTEIN TYROSINE PHOSPHATASE PTP20
; AND RELATED PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,626A
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/951,260
; FILING DATE: October 16, 1997
; APPLICATION NUMBER: 60/030,860
; FILING DATE: No. 6482605ember 13, 1996
; APPLICATION NUMBER: PCT/1897/00946
; FILING DATE: June 17, 1997
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: "Xaa" in positions 3 and 5 stand
; for an unspecified amino acid.
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-430-626A-2

Query Match 73.7%; Score 28; DB 4; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
DB 1 FWXXW 6

RESULT 8
US-09-361-096A-46
; Sequence 46, Application US/09361096A
; Patent No. 6492495
; GENERAL INFORMATION:
; APPLICANT: MOLLER, NIELS P.H.
; APPLICANT: MOLLER, KARIN B.
; APPLICANT: ULLRICH, AXEL
; TITLE OF INVENTION: PTP-531: A NOVEL PROTEIN TYROSINE PHOSPHATASE
; FILE REFERENCE: 038602/0686
; CURRENT APPLICATION NUMBER: US/09/361,096A
; CURRENT FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 08/449,609
; PRIOR FILING DATE: 1995-05-24
; PRIOR APPLICATION NUMBER: 08/036,210
; PRIOR FILING DATE: 1995-03-23
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Consensus
; OTHER INFORMATION: sequence
; NAME/KEY: MOD_RES
; LOCATION: (3)
; OTHER INFORMATION: Variable amino acid
; NAME/KEY: MOD_RES
; LOCATION: (5)
; OTHER INFORMATION: Variable amino acid
US-09-361-096A-46

Query Match 73.7%; Score 28; DB 4; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
DB 1 FWXXW 6

RESULT 9
US-09-743-492A-5
; Sequence 5, Application US/09743492A
; Patent No. 6709843
; GENERAL INFORMATION:
; APPLICANT: YAMAMOTO, Hiroshi
; APPLICANT: TSUKIKAWA, Kazutake
; APPLICANT: UCHINO, Yukiko
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR INTRACELLULAR DOMAIN OF PROTEIN
; FILE REFERENCE: 19036/37023
; CURRENT APPLICATION NUMBER: US/09/743,492A
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: PCT/JP98/03120
; PRIOR FILING DATE: 1998-07-10
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Consensus Amino
; OTHER INFORMATION: Acid Sequence in Cytoplasmic Domain of Known PTPs.
; Patent No. 6709843
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa= Arg, Glu or Leu
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa= Val, Ile or Cys
US-09-743-492A-5

Query Match 73.7%; Score 28; DB 4; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
DB 1 FWXXW 6

RESULT 10
US-10-243-687-2
; Sequence 2, Application US/10243687
; Patent No. 6797501
; GENERAL INFORMATION:
; APPLICANT: Aoki, Naohito
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PROTEIN TYROSINE PHOSPHATASE PTP20
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/243,687
; FILING DATE: 16-Sep-2002
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,626A
; FILING DATE: 29-Oct-1999
```


APPLICATION NUMBER: 08/951,260
FILING DATE: October 16, 1997
APPLICATION NUMBER: 60/030,860
FILING DATE: No. 6797501ember 13, 1996
APPLICATION NUMBER: PCT/1897/00946
FILING DATE: June 17, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: "Xaa" in positions 3 and 5 stand
for an unspecified amino acid.
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-243-687-2

Query Match 73.7%; Score 28; DB 4; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
DB 1 FWMXXW 6

RESULT 11
US-07-973-235A-30
Sequence 30, Application US/07973235A
Patent No. 5491130
GENERAL INFORMATION:
APPLICANT: David D. Roberts, et al.
TITLE OF INVENTION: Peptide Inhibitors of Fibronectin and
RELATED COLLAGEN-BINDING PROTEINS
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lowe, Price, LeBlanc & Becker
STREET: Suite 300, 99 Canal Center Plaza
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect, Version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/973,235A
FILING DATE: 19921110
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Robert L. Price
REGISTRATION NUMBER: 22,685
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)684-1111
TELEFAX: (703)684-1124
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: AMINO ACIDS
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-07-973-235A-30

Query Match 73.7%; Score 28; DB 1; Length 7;
Best Local Similarity 42.9%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXXW 7
DB 1 GWSKSW 7

RESULT 12
US-08-652-971-5
Sequence 5, Application US/08652971
Patent No. 5814507
GENERAL INFORMATION:
APPLICANT: Cheng, Jill
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
PHOSPHATASE, PTP LAMBDA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd.
CITY: South San Francisco
STATE: California
COUNTRY: United States
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,971
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P1033
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 225-3216
TELEFAX: (415) 952-9881
TELEX: 910 371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Active-site
LOCATION: 1..2
OTHER INFORMATION: /note= "Let 'X' located at position
OTHER INFORMATION: 1 represent either Histidine or Aspartic Acid"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6..7
OTHER INFORMATION: /note= "Let 'X' located at position
OTHER INFORMATION: 6 represent either Isoleucine or Valine."
US-08-652-971-5

Query Match 73.7%; Score 28; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
DB 2 FWMXXW 7

RESULT 13
US-08-462-720-30
; Sequence 30, Application US/08462720
; Patent No. 5849701
; GENERAL INFORMATION:
; APPLICANT: Roberts, David D.
; APPLICANT: Kruttsch, Henry C.
; APPLICANT: Sipes, John M.
; APPLICANT: Guo, Neng-hua
; APPLICANT: Negre, Eric
; TITLE OF INVENTION: Peptide Inhibitors of Fibronectin and
; Related Collagen-Binding Proteins
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,720
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 015280-023110US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-462-720-30

Query Match 73.7%; Score 28; DB 2; Length 7;
Best Local Similarity 42.9%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GWXXW 7
| |::|
Db 1 GWSKSW 7

RESULT 14
US-08-991-258A-5
; Sequence 5, Application US/08991258A
; Patent No. 592887
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; APPLICANT: Lasky, Laurence A.
; TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
; PHOSPHATASE, PTP LAMBDA
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOEBACH, TEST, ALBERTON & HERBERT, LLP
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,258A
; FILING DATE: 17-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,971
; FILING DATE: 24-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-63478-3/WH/D/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Active-site
; LOCATION: 1..2
; OTHER INFORMATION: /note= "Let 'X' located at position
; OTHER INFORMATION: 1 represent either Histidine or Aspartic Acid"
US-08-991-258A-5

Query Match 73.7%; Score 28; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
:|::|
Db 2 FWRXW 7

RESULT 15
US-08-769-399-5
; Sequence 5, Application US/08769399
; Patent No. 5976852
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; APPLICANT: Lasky, Laurence A.
; TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
; PHOSPHATASE, PTP LAMBDA
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd.
; CITY: South San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,399
; FILING DATE:
; CLASSIFICATION: 435

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Dregger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 225-3216
; TELEFAX: (415) 952-9881
; TELEX: 910 371-7168
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Active-site
; LOCATION: 1..2
; OTHER INFORMATION: /note= "Let 'X' located at position
; OTHER INFORMATION: 1 represent either Histidine or Aspartic Acid"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 6..7
; OTHER INFORMATION: /note= "Let 'X' located at position
; OTHER INFORMATION: 6 represent either Isoleucine or Valine."
US-08-769-399-5

Query Match 73.7%; Score 28; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
Db :|||
2 FWRMW 7

Search completed: January 3, 2005, 17:27:14
Job time : 37 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 17:15:18 ; Search time 139 Seconds
(without alignments)
18.116 Million cell updates/sec

Title: US-10-046-922-67

Perfect score: 38

Sequence: 1 GYXXXXX 7

Scoring table: BLOSUM62DX-1

Searched: Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 60586

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	7	13	US-10-046-922-67
2	32	84.2	6	15	US-10-418-943-48
3	28	73.7	6	13	US-10-087-993-1
4	28	73.7	6	14	US-10-243-687-2
5	28	73.7	6	14	US-10-314-232-46
6	28	73.7	7	17	US-10-858-271-10
7	27	71.1	7	9	US-09-884-767A-38
8	27	71.1	7	10	US-09-563-222-53
9	27	71.1	7	14	US-10-052-578-183
10	27	71.1	7	14	US-10-052-578-195
11	27	71.1	7	14	US-10-190-082-65
12	27	71.1	7	14	US-10-190-082-76
13	27	71.1	7	14	US-10-053-520-183

14	27	71.1	7	14	US-10-053-520-195	Sequence 195, App
15	27	71.1	7	14	US-10-053-498B-183	Sequence 183, App
16	27	71.1	7	14	US-10-053-498B-195	Sequence 195, App
17	27	71.1	7	15	US-10-418-943-71	Sequence 71, Appl
18	27	71.1	7	15	US-10-258-146A-41	Sequence 41, Appl
19	27	71.1	7	15	US-10-258-146A-53	Sequence 53, Appl
20	27	71.1	7	15	US-10-328-953-186	Sequence 186, App
21	27	71.1	7	15	US-10-328-953-198	Sequence 198, App
22	27	71.1	7	16	US-10-258-144-76	Sequence 76, Appl
23	27	71.1	7	16	US-10-258-144-88	Sequence 88, Appl
24	27	71.1	7	16	US-10-475-853-6	Sequence 6, Appl
25	27	71.1	7	17	US-10-783-950-53	Sequence 53, Appl
26	26	68.4	5	15	US-10-436-549-496	Sequence 496, App
27	26	68.4	5	16	US-10-712-425-496	Sequence 496, App
28	26	68.4	7	15	US-10-403-938-27	Sequence 27, Appl
29	25	65.8	5	10	US-09-750-754-3	Sequence 3, Appl
30	25	65.8	5	10	US-09-750-754-15	Sequence 15, Appl
31	25	65.8	5	10	US-09-915-914B-18	Sequence 18, Appl
32	25	65.8	5	10	US-09-962-756-1586	Sequence 1586, Ap
33	25	65.8	5	14	US-10-253-471-1586	Sequence 1586, Ap
34	25	65.8	5	15	US-10-253-493-1586	Sequence 1586, Ap
35	25	65.8	5	15	US-10-436-549-101	Sequence 101, App
36	25	65.8	5	15	US-10-436-549-525	Sequence 525, App
37	25	65.8	5	15	US-10-243-613-84	Sequence 84, Appl
38	25	65.8	5	15	US-10-664-021-69	Sequence 69, Appl
39	25	65.8	5	15	US-10-374-466-14	Sequence 14, Appl
40	25	65.8	5	16	US-10-712-425-101	Sequence 101, App
41	25	65.8	5	16	US-10-712-425-525	Sequence 525, App
42	25	65.8	5	17	US-10-654-578-2007	Sequence 2007, Ap
43	25	65.8	6	9	US-09-766-396-19	Sequence 19, Appl
44	25	65.8	6	9	US-09-788-268-2	Sequence 2, Appl
45	25	65.8	6	13	US-10-062-375-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1

US-10-046-922-67
; Sequence 67, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 289677/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 67
; LENGTH: 7
; TYPE: PRT
; ORGANISM: peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)..(6)
; OTHER INFORMATION: X at position 4-6 is any amino acid
US-10-046-922-67

Query Match 100.0%; Score 38; DB 13; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXX 7
| | | | | | |
Db 1 GYXXXXX 7

RESULT 2

US-10-418-943-48
; Sequence 48, Application US/10418943

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:

OTHER INFORMATION: "Xaa" in positions 3 and 5 stand
for an unspecified amino acid.
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-243-687-2

Query Match 73.7%; Score 28; DB 14; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.5e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
:|:|:
Db 1 FWXXW 6

RESULT 5

US-10-314-232-46

; Sequence 46, Application US/10314232
; Publication No. US20030138932A1

GENERAL INFORMATION:

APPLICANT: MOLLER, NIELS P.H.
APPLICANT: MOLLER, KARIN B.
APPLICANT: ULLRICH, AXEL

TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE PHOSPHATASE

FILE REFERENCE: 038602/0686

CURRENT APPLICATION NUMBER: US/10/314,232

CURRENT FILING DATE: 2002-12-09

PRIOR APPLICATION NUMBER: US/09/361,096

PRIOR FILING DATE: 1999-07-26

PRIOR APPLICATION NUMBER: 08/449,609

PRIOR FILING DATE: 1995-05-24

PRIOR APPLICATION NUMBER: 08/036,210

PRIOR FILING DATE: 1995-03-23

NUMBER OF SEQ ID NOS: 53

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 46

LENGTH: 6

TYPE: PRT

ORGANISM: Artificial Sequence

OTHER INFORMATION: Description of Artificial Sequence: Consensus
FEATURE:

OTHER INFORMATION: sequence

FEATURE:

LOCATION: (3)_

NAME/KEY: MOD_RES

FEATURE:

OTHER INFORMATION: Variable amino acid

NAME/KEY: MOD_RES

LOCATION: (5)

OTHER INFORMATION: Variable amino acid

US-10-314-232-46

Query Match 73.7%; Score 28; DB 14; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.5e+06;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
:|:|:
Db 1 FWXXW 6

RESULT 6

US-10-858-271-10

; Sequence 10, Application US/10858271
; Publication No. US20040259829A1

GENERAL INFORMATION:

APPLICANT: Danks, Mary K.

APPLICANT: Potter, Philip M.

APPLICANT: Houghton, Peter J.

TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of

TITLE OF INVENTION: Tumor Cells

FILE REFERENCE: SJ-0005

CURRENT APPLICATION NUMBER: US/10/858,271

CURRENT FILING DATE: 2004-06-01

PRIOR APPLICATION NUMBER: US/09/595,682

PRIOR FILING DATE: 2000-06-16

PRIOR APPLICATION NUMBER: 60/075,258

PRIOR FILING DATE: 1998-02-19

PRIOR APPLICATION NUMBER: PCT/US99/03171

PRIOR FILING DATE: 1999-02-12

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 10

LENGTH: 7

TYPE: PRT

ORGANISM: Oryctolagus cuniculus

US-10-858-271-10

Query Match 73.7%; Score 28; DB 17; Length 7;
Best Local Similarity 33.3%; Pred. No. 1.5e+06;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
:|:|:
Db 2 FWXXW 7

RESULT 7

US-09-884-767A-38

Sequence 38, Application US/09884767A

Publication No. US20020192789A1

GENERAL INFORMATION:

APPLICANT: DYAX Corp.

APPLICANT: Ley, Arthur C.

APPLICANT: Luneau, Christopher J.

APPLICANT: Ladner, Robert C

TITLE OF INVENTION: NOVEL ENTEROKINASE CLEAVAGE SEQUENCES

FILE REFERENCE: DYX-012.1 US, DYX-012.1 PCT

CURRENT APPLICATION NUMBER: US/09/884,767A

CURRENT FILING DATE: 2001-06-19

PRIOR APPLICATION NUMBER: US 09/597,321

PRIOR FILING DATE: 2000-06-19

NUMBER OF SEQ ID NOS: 217

SOFTWARE: Patentin version 3.1

SEQ ID NO 38

LENGTH: 7

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: synthetic enterokinase cleavage sequence
US-09-884-767A-38

Query Match 71.1%; Score 27; DB 9; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWXXX 6
:|:|:
Db 1 GYWDR 6

RESULT 8

US-09-563-222-53

Sequence 53, Application US/09563222

Publication No. US20030079253A1

; GENERAL INFORMATION:
; APPLICANT: Hiatt, Andrew
; APPLICANT: Hein, Mich B.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN
; FILE REFERENCE: 310098.406
; CURRENT APPLICATION NUMBER: US/09/563,222
; CURRENT FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-563-222-53

Query Match 71.1%; Score 27; DB 10; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXX 6
|||:::
Db 2 GYWNNS 7

RESULT 9

US-10-052-578-183
; Sequence 183, Application US/10052578
; Publication No. US20030134787A1
; GENERAL INFORMATION:
; APPLICANT: Sloan-Kettering Institute for Cancer Research
; APPLICANT: Rothman, James E.
; APPLICANT: Mayhew, Mark
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Hartl, Ulrich
; APPLICANT: Querfelli, Ouathak
; APPLICANT: Morol, Yoichi
; TITLE OF INVENTION: CONJUGATE HEAT SHOCK PROTEIN-BINDING PEPTIDES
; FILE REFERENCE: 11746/46003
; CURRENT APPLICATION NUMBER: US/10/052,578
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 08/961,707
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 183
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide in m13 coliphage
US-10-052-578-183

Query Match 71.1%; Score 27; DB 14; Length 7;
Best Local Similarity 33.3%; Pred. No. 1.5e+06;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
:|::|
Db 2 HWWDW 7

RESULT 10

US-10-052-578-195
; Sequence 195, Application US/10052578
; Publication No. US20030134787A1
; GENERAL INFORMATION:
; APPLICANT: Sloan-Kettering Institute for Cancer Research
; APPLICANT: Rothman, James E.
; APPLICANT: Mayhew, Mark
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan

; APPLICANT: Hartl, Ulrich
; APPLICANT: Querfelli, Ouathak
; APPLICANT: Morol, Yoichi
; TITLE OF INVENTION: CONJUGATE HEAT SHOCK PROTEIN-BINDING PEPTIDES
; FILE REFERENCE: 11746/46003
; CURRENT APPLICATION NUMBER: US/10/052,578
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 08/961,707
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 195
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide in m13 coliphage
US-10-052-578-195

Query Match 71.1%; Score 27; DB 14; Length 7;
Best Local Similarity 33.3%; Pred. No. 1.5e+06;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
:|::|
Db 2 HWQMLW 7

RESULT 11

US-10-190-082-65
; Sequence 65, Application US/10190082
; Publication No. US20030148264A1
; GENERAL INFORMATION:
; APPLICANT: Lasky, Lawrence A.
; APPLICANT: Sidhu, Sachdev S.
; APPLICANT: Held, Heike A.
; TITLE OF INVENTION: PHAGE DISPLAYED PDZ DOMAIN LIGANDS
; FILE REFERENCE: P1905R1
; CURRENT APPLICATION NUMBER: US/10/190,082
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,634
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 683
; SEQ ID NO 65
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-190-082-65

Query Match 71.1%; Score 27; DB 14; Length 7;
Best Local Similarity 33.3%; Pred. No. 1.5e+06;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
:|::|
Db 1 WWADV 6

RESULT 12

US-10-190-082-76
; Sequence 76, Application US/10190082
; Publication No. US20030148264A1
; GENERAL INFORMATION:
; APPLICANT: Lasky, Lawrence A.
; APPLICANT: Sidhu, Sachdev S.
; APPLICANT: Held, Heike A.
; TITLE OF INVENTION: PHAGE DISPLAYED PDZ DOMAIN LIGANDS
; FILE REFERENCE: P1905R1
; CURRENT APPLICATION NUMBER: US/10/190,082
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,634

; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 683
; SEQ ID NO 76
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-190-082-76

Query Match 71.1%; Score 27; DB 14; Length 7;
Best Local Similarity 33.3%; Pred. No. 1.5e+06;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7
:|::|
DB 1 HWIETW 6

RESULT 13

US-10-053-520-183
; Sequence 183, Application US/10053520
; Publication No. US20030166530A1
; GENERAL INFORMATION:
; APPLICANT: Sloan-Kettering Institute for Cancer Research
; APPLICANT: Rothman, James E.
; APPLICANT: Mayhew, Mark
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Hartl, Ulrich
; APPLICANT: Querfelli, Ouathek
; APPLICANT: Moroi, Yoichi
; TITLE OF INVENTION: CONJUGATE HEAT SHOCK PROTEIN-BINDING PEPTIDES
; FILE REFERENCE: 11746/46004
; CURRENT APPLICATION NUMBER: US/10/053,520
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 08/961,707
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 183
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide in m13 coliphage
US-10-053-520-183

Query Match 71.1%; Score 27; DB 14; Length 7;
Best Local Similarity 33.3%; Pred. No. 1.5e+06;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7
:|::|
DB 2 HWDWW 7

RESULT 14

US-10-053-520-195
; Sequence 195, Application US/10053520
; Publication No. US20030166530A1
; GENERAL INFORMATION:
; APPLICANT: Sloan-Kettering Institute for Cancer Research
; APPLICANT: Rothman, James E.
; APPLICANT: Mayhew, Mark
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Hartl, Ulrich
; APPLICANT: Querfelli, Ouathek
; APPLICANT: Moroi, Yoichi
; TITLE OF INVENTION: CONJUGATE HEAT SHOCK PROTEIN-BINDING PEPTIDES
; FILE REFERENCE: 11746/46004
; CURRENT APPLICATION NUMBER: US/10/053,520

; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 08/961,707
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 195
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide in m13 coliphage
US-10-053-520-195

Query Match 71.1%; Score 27; DB 14; Length 7;
Best Local Similarity 33.3%; Pred. No. 1.5e+06;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7
:|::|
DB 2 HWQMLW 7

RESULT 15

US-10-053-498B-183
; Sequence 183, Application US/10053498B
; Publication No. US20030194409A1
; GENERAL INFORMATION:
; APPLICANT: Sloan-Kettering Institute for Cancer Research
; APPLICANT: Rothman, James E.
; APPLICANT: Mayhew, Mark
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Hartl, Ulrich
; APPLICANT: Querfelli, Ouathek
; APPLICANT: Moroi, Yoichi
; TITLE OF INVENTION: CONJUGATE HEAT SHOCK PROTEIN-BINDING PEPTIDES
; FILE REFERENCE: 11746/46002
; CURRENT APPLICATION NUMBER: US/10/053,498B
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 08/961,707
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 183
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide in m13 coliphage
US-10-053-498B-183

Query Match 71.1%; Score 27; DB 14; Length 7;
Best Local Similarity 33.3%; Pred. No. 1.5e+06;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7
:|::|
DB 2 HWDWW 7

Search completed: January 3, 2005, 17:29:39
Job time : 140 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 17:01:32 ; Search time 38 Seconds
(without alignments)
17.724 Million cell updates/sec

Title: US-10-046-922-67

Perfect score: 38

Sequence: 1 GYWXW 7

Scoring table: BLOSUM62X
Gapop 10.0 ; Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 457

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	57.9	5	2	JH0253	gut pentapeptide -
2	17	44.7	6	2	B35640	cerebellar degener
3	17	44.7	7	2	S21230	dermorphin (trp-4,
4	16	42.1	6	2	F41946	T-cell receptor ga
5	16	42.1	7	2	PT0728	T-cell receptor be
6	15	39.5	6	2	PT0629	T-cell receptor be
7	15	39.5	6	2	PT0532	T-cell receptor be
8	15	39.5	7	2	PN0649	pullulanase (EC 3.
9	14	36.8	4	2	A34626	RPCH-related neuro
10	14	36.8	5	2	A32516	cholecystokinin-5
11	14	36.8	5	2	A60803	neuropeptide - sea
12	14	36.8	5	2	PT0281	Ig heavy chain CRD
13	14	36.8	5	2	PT0308	Ig heavy chain CRD
14	14	36.8	5	2	PT0729	T-cell receptor be
15	14	36.8	5	2	PT0580	T-cell receptor be
16	14	36.8	6	2	S66195	alcohol dehydrogen
17	14	36.8	6	2	B34835	dnaA protein - Pse
18	14	36.8	6	2	A31263	dihydrofolate redu
19	14	36.8	6	2	B31263	T-cell receptor be
20	14	36.8	6	2	PT0519	T-cell receptor be
21	14	36.8	6	2	PT0637	T-cell receptor be
22	14	36.8	6	2	PT0641	T-cell receptor be
23	14	36.8	6	2	PT0726	T-cell receptor be
24	14	36.8	6	2	PD0028	pev-kinin 2 - pena
25	14	36.8	6	2	A61068	locustakinin -mig
26	14	36.8	6	2	I79564	hypothetical TCL3
27	14	36.8	7	2	S09652	hypothetical prote
28	14	36.8	7	2	PQ0727	H2 class I protein
29	14	36.8	7	2	E48394	glycoprotein compo

30 14 36.8 7 2 PH1602 Ig H chain V-D-J r
31 14 36.8 7 2 PT0526 T-cell receptor be
32 14 36.8 7 2 PT0628 T-cell receptor be
33 14 36.8 7 2 PT0642 T-cell receptor be
34 14 36.8 7 2 PT0722 T-cell receptor be
35 14 36.8 7 2 PT0688 T-cell receptor be
36 14 36.8 7 2 PT0586 T-cell receptor be
37 14 36.8 7 2 PX0008 glucuronosyltransf
38 14 36.8 7 2 B48194 major fat-globule
39 14 36.8 7 2 PD0029 pev-kinin 1 - pena
40 14 36.8 7 2 S57274 triacylglycerol 11
41 14 36.8 7 2 S33244 neuromodulatory pe
42 14 36.8 7 2 S33245 neuromodulatory pe
43 14 36.8 7 2 S33246 neuromodulatory pe
44 14 36.8 7 2 S33567 tubulin beta-3 cha
45 14 36.8 7 2 A58512 venom heptapeptide

ALIGNMENTS

RESULT 1

JH0253

gut pentapeptide - Japanese eel

C:Species: Anguilla japonica (Japanese eel)

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Apr-1995

C:Accession: JH0253

R:Uesaka, T.; Ikeda, T.; Kubota, I.; Muneoka, Y.; Ando, M.

Biochem. Biophys. Res. Commun. 180, 828-832, 1991

A:Title: Structure and function of a pentapeptide isolated from the gut of the eel.

A:Reference number: JH0253; MUID:92062113; PMID:1953755

A:Accession: JH0253

A:Molecule type: protein

A:Residues: 1-5 <UES>

A:Experimental source: gut

C:Comment: This peptide increased basal tone of the circular muscle of the esophagogastric

, and of the circular muscle of the gastro-intestinal junction.

Query Match 57.9%; Score 22; DB 2; Length 5;

Best Local Similarity 40.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWX 5

Db 1 GFWK 5

RESULT 2

B35640

cerebellar degeneration-related protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 24-Jun-1993

C:Accession: B35640

R:Chen, Y.T.; Rettig, W.J.; Yenamandra, A.K.; Kozak, C.A.; Chaganti, R.S.K.; Posner, J.B

Proc. Natl. Acad. Sci. U.S.A. 87, 3077-3081, 1990

A:Title: Cerebellar degeneration-related antigen: a highly conserved neuroectodermal mar

A:Reference number: A35640; MUID:90222173; PMID:2326268

A:Accession: B35640

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-6 <CHE>

Query Match 44.7%; Score 17; DB 2; Length 6;

Best Local Similarity 20.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXX 6

Db 1 FWEDL 5

RESULT 3

S21230

```

dermorphin (Trp-4, Asn-7) [validated] - two-colored leaf frog (fragment)
C:Species: Phyllomedusa bicolor (two-colored leaf frog)
C:Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 18-Aug-2000
C:Accession: S21230
R:Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erapamer, G.; Kreil, G.
FBS Lett. 302, 151-154, 1992
A:Title: Identification and characterization of two dermorphins from skin extracts of th
A:Reference number: S21152; MUID:92339502; PMID:1633846
A:Accession: S21230
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <MIG>
C:Superfamily: dermorphin precursor; dermorphin precursor amino-terminal homology

Query Match      44.7%; Score 17; DB 2; Length 7;
Best Local Similarity 20.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXX 6
   :|::
Db 3 FWYPN 7

RESULT 4
F41946
T-cell receptor gamma chain (1a.27) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C:Accession: F41946
R:WhetSELL, M.; Mosley, R.L.; WhetSELL, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
A:Reference number: A41946; MUID:92049316; PMID:1658619
A:Accession: F41946
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-6 <WHE>
C:Keywords: T-cell receptor

Query Match      42.1%; Score 16; DB 2; Length 6;
Best Local Similarity 20.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXW 7
   :|::|
Db 1 YCAVV 5

RESULT 5
PT0728
T-cell receptor beta chain V-D-J region (161-2H) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0728
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0728
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-7 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match      42.1%; Score 16; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXX 5
   |::
Db 3 GDWGW 7

```

```

RESULT 6
PT0629
T-cell receptor beta chain V-D-J region (100-2AH) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0629; PT0528
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0629
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FEE>
A:Experimental source: newborn thymus, strain BALB/c, clone 100-2AH
A:Accession: PT0528
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FE2>
A:Experimental source: adult thymus, strain BALB/c, clone 100-4AB
C:Keywords: T-cell receptor

Query Match      39.5%; Score 15; DB 2; Length 6;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWX 4
   |::
Db 3 GDWG 6

RESULT 7
PT0532
T-cell receptor beta chain V-D-J region (100-4AJ) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0532
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0532
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FEE>
A:Experimental source: adult thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match      39.5%; Score 15; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYW 3
   |::
Db 4 GYW 6

RESULT 8
PN0649
pullulanase (EC 3.2.1.41) - Bacillus sp. (strain S-1) (fragment)
C:Species: Bacillus sp.
C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 03-Jun-2002
C:Accession: PN0649
R:Kim, C.H.; Choi, H.I.; Lee, D.S.
Biosci. Biotechnol. Biochem. 57, 1632-1637, 1993
A:Title: Purification and biochemical properties of an alkaline pullulanase from alkalop
A:Reference number: PN0649; MUID:94080025; PMID:7764261
A:Accession: PN0649
A:Molecule type: protein
A:Residues: 1-7 <KIM>
C:Comment: This enzyme is used together with glucoamylase to improve the efficiency of st
ntent in high maltose syrups.

```

C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 39.5%; Score 15; DB 2; Length 7;
Best Local Similarity 16.7%; Pred. No. 2.8e+05;
Matches 1; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 YWXXW 7
DB 1 FLNMSW 6

RESULT 9

A34626
RPCH-related neuropeptide - ferruginous spindle
C;Species: *Fusinus ferrugineus* (ferruginous spindle)
C;Date: 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change 31-Dec-1993
C;Accession: A34626
R;Kuroki, Y.; Kanda, T.; Kubota, I.; Fujisawa, Y.; Ikeda, T.; Miura, A.; Minamitake, Y.;
Biochem. Biophys. Res. Commun. 167, 273-278, 1990
A;Title: A molluscan neuropeptide related to the crustacean hormone, RPCH.
A;Reference number: A34626; MUID:90179762; PMID:2310394
A;Accession: A34626
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-4 <KUR>
C;Keywords: neuropeptide

Query Match 36.8%; Score 14; DB 2; Length 4;
Best Local Similarity 25.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 XXXW 7
DB 1 AFGW 4

RESULT 10

A32516
cholecystokinin-5 - dog
N;Alternate names: CCK-5
C;Species: *Canis lupus familiaris* (dog)
C;Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 18-Aug-2000
C;Accession: A32516
R;Shively, J.; Reeve Jr., J.R.; Byssselein, V.E.; Ben-Avram, C.; Vigna, S.R.; Walsh, J.H.;
Am. J. Physiol. 252, G272-G275, 1987
A;Title: CCK-5: sequence analysis of a small cholecystokinin from canine brain and intest
A;Reference number: A32516; MUID:87153871; PMID:3826354
A;Accession: A32516
A;Molecule type: protein
A;Residues: 1-5 <SHI>
C;Comment: This peptide corresponds to the five carboxyl-terminal residues of cholecysto
C;Superfamily: gastrin
C;Keywords: amidated carboxyl end; neuropeptide
F;5/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 36.8%; Score 14; DB 2; Length 5;
Best Local Similarity 25.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXX 6
DB 2 WMDP 5

RESULT 11

A60803
neuropeptide - sea anemone (*Anthopleura elegantissima*)
C;Species: *Anthopleura elegantissima*
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: A60803
R;Graff, D.; Grimmelikhuijzen, C.J.P.
Brain Res. 442, 354-358, 1988
A;Title: Isolation of <Glu-Ser-Lu-Arg-Trp-NH-2, a novel neuropeptide from sea anemones.

A;Reference number: A60803; MUID:88222764; PMID:2897223

A;Accession: A60803
A;Molecule type: protein
A;Residues: 1-5 <GRA>
C;Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid
F;1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F;5/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 36.8%; Score 14; DB 2; Length 5;
Best Local Similarity 25.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 XXXW 7
DB 2 SLRW 5

RESULT 12

PT0281
Ig heavy chain CRD3 region (clone 4-91C) - human (fragment)
C;Species: *Homo sapiens* (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0281
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0281
A;Molecule type: DNA
A;Residues: 1-5 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 36.8%; Score 14; DB 2; Length 5;
Best Local Similarity 25.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 XXXW 7
DB 1 DENW 4

RESULT 13

PT0308
Ig heavy chain CRD3 region (clone 6-88) - human (fragment)
C;Species: *Homo sapiens* (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0308
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0308
A;Molecule type: DNA
A;Residues: 1-5 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 36.8%; Score 14; DB 2; Length 5;
Best Local Similarity 25.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXX 6
DB 2 WESS 5

RESULT 14

PT0729
T-cell receptor beta chain V-D-J region (120-1J) - mouse (fragment)
C;Species: *Mus musculus* (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0640; PT0685; PT0729

R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0640
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: newborn thymus, strain BALB/c, clone 120-1J
A;Accession: PT0685
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-5 <FE2>
A;Experimental source: day 18 fetal thymus, strain BALB/c, clone 154-1C
A;Accession: PT0729
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-5 <FE3>
A;Experimental source: newborn thymus, strain BALB/c, clone 135-1AG
C;Keywords: T-cell receptor

Query Match 36.8%; Score 14; DB 2; Length 5;
Best Local Similarity 25.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 XXXW 7
:::
Db 2 SSDW 5

RESULT 15

PT0580
T-cell receptor beta chain V-D-J region (159-2B) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0580

R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0580
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 36.8%; Score 14; DB 2; Length 5;
Best Local Similarity 25.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 XXXW 7
:::
Db 1 ASSW 4

Search completed: January 3, 2005, 17:15:13
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 17:09:38 ; Search time 185 Seconds
(without alignments)
21.771 Million cell updates/sec

Title: US-10-046-922-67

Perfect score: 38

Sequence: 1 GYXXW 7

Scoring table: BLOSUM62, Gapop 10.0, Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 167

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: UniProt_02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	65.8	5	1	UF01_MOUSE
2	16	42.1	4	1	OCP3_OCTMI
3	16	42.1	7	2	Q9BRV4
4	15	39.5	6	1	E101_LITRU
5	14	36.8	6	1	LOK1_LOCM1
6	14	36.8	7	1	BRHP_CONTM
7	14	36.8	7	1	TPFY_PACDA
8	14	36.8	7	1	TY51_LITRU
9	14	36.8	7	1	WMA1_ACHFU
10	14	36.8	7	1	WMA2_ACHFU
11	14	36.8	7	1	WMA3_ACHFU
12	14	36.8	7	2	Q95945
13	14	36.8	7	2	O49223
14	14	36.8	7	2	Q8GL00
15	14	36.8	7	2	Q8GL04
16	14	36.8	7	2	Q8KMS9
17	14	36.8	7	2	Q8JEB1
18	13	34.2	5	1	BPPT_BOTIN
19	12	31.6	2	1	GWA_SEPOF
20	12	31.6	5	1	PAP2_PARMA
21	11	28.9	5	1	RE32_LITRU
22	9	23.7	4	1	ACH1_ACHFU
23	9	23.7	4	1	OCPI_OCTMI
24	9	23.7	7	1	FARS_HIRME
25	9	23.7	7	1	UN06_PINPS
26	8	21.1	3	1	GRWM_HUMAN
27	8	21.1	4	1	DCML_PSECH
28	8	21.1	4	1	PAR4_HIRME
29	8	21.1	6	1	FARP_MONEX
30	8	21.1	7	1	ASCL_ALLAS
31	8	21.1	7	1	FAR3_HAECO

32	8	21.1	7	1	FAR3_PANRE	P41874 panagrellus
33	8	21.1	7	1	GFRP_MOUSE	P99025 mus musculus
34	8	21.1	7	1	HV7_FIG	P01153 sus scrofa
35	8	21.1	7	2	O99182	O99182 gnatholebia
36	8	21.1	7	2	P82445	P82445 nicotiana c
37	8	21.1	7	2	Q9YQ10	Q9YQ10 transmissib
38	7	18.4	4	1	FAR3_HIRME	P42562 hirudo medi
39	7	18.4	4	1	FYR1_ANTEL	P58706 anthopleura
40	7	18.4	4	2	O96AT0	O96AT0 homo sapien
41	7	18.4	5	1	AL14_CARMA	P81817 carinus ma
42	7	18.4	5	1	FARP_ARTTR	P41853 artiposthi
43	7	18.4	5	1	PRCT_PERAM	P01373 periplaneta
44	7	18.4	5	1	PSK_DAUCA	P58261 daucus caro
45	7	18.4	6	1	ASP2_LACSN	P82655 lactobacill

ALIGNMENTS

RESULT 1
UF01_MOUSE
ID UF01_MOUSE STANDARD; PRT; 5 AA.
AC P38639;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Unknown protein from 2D-PAGE of fibroblasts (P19) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Fibroblast;
RX MEDLINE=95009907; PubMed=7523108;
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT "Separation and sequencing of familial and novel murine proteins using
RT Preparative two-dimensional gel electrophoresis";
RL Electrophoresis 15:735-745(1994).
CC -1- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC protein is: 6.6, its MW is: 19 kDa.
KW Direct protein sequencing.
FT NON_TER 5
SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;

Query Match 65.8%; Score 25; DB 1; Length 5;
Best Local Similarity 40.0%; Pred. No. 1.8e+06;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXW 7
|:|:
Db 1 WIGRW 5

RESULT 2
OCPI_OCTMI
ID OCP3_OCTMI STANDARD; PRT; 4 AA.
AC P5849;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Cardioactive peptides Ocp-3/Ocp-4.
OS Octopus minor (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
RN [1]
RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
RC TISSUE=Brain;
RX MEDLINE=20336815; PubMed=10876044;
RA Iwakoshi E., Hisada M., Minakata H.;
RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
RT Octopus minor.";

RL Peptides 21:623-630(2000).
CC -I- FUNCTION: Cardioactive; has both positive chronotropic and
CC inotropic effects on the heart. Ocp-4 is a 1000 time less active
CC than Ocp-3.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- PTM: Ocp-4 has D-Ser instead of L-Ser.
CC -I- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI; RANGE=1-4; NOTE=Ref.1.
CC -I- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI; RANGE=1-4; NOTE=Ref.1.
KW D-amino acid; Direct protein sequencing; Hormone.
FT MOD RES 2 2 D-serine (in form Ocp-4).
SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;

Query Match 42.1%; Score 16; DB 1; Length 4;
Best Local Similarity 50.0%; Pred. No. 1.8e+06;
Matches 2; Conservative 1; Mismatches 0; Gaps 0;

QY 1 GYWX 4
DB 1 GSWD 4

RESULT 3
Q9BRY4 PRELIMINARY; PRT; 7 AA.
ID Q9BRY4
AC Q9BRY4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SQSTM1 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Hopkins R.F., Jordan B., Buetow K.H., Schaefer C.F., Bhat N.K., Altschul S.F., Zerbahn T., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Diatchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005857; AAH05857.3; -;
DR GO; GO:0005829; C:cytosol; ISS.
DR GO; GO:0019901; F:protein kinase binding; ISS.
DR GO; GO:0042169; F:SH2 domain binding; ISS.
DR GO; GO:0043130; F:ubiquitin binding; ISS.
DR GO; GO:0016197; P:endosome transport; ISS.
DR GO; GO:0007242; P:intracellular signaling cascade; ISS.
DR GO; GO:0045944; P:positive regulation of transcription from P. . .; ISS.
DR GO; GO:0008104; P:protein localization; ISS.
DR GO; GO:0043122; P:regulation of I-kappaB kinase/NF-kappaB cas. . .; ISS.
DR GO; GO:0006950; P:response to stress; ISS.
DR InterPro; IPR000449; UBA.

DR InterPro; IPR000433; Znf_ZZ.
DR Pfam; PF00569; ZZ; 1.
DR SMART; SM00165; UBA; 1.
DR SMART; SM00291; Znf_ZZ; 1.
DR PROSITE; PS00030; UBA; 1.
DR PROSITE; PS01357; ZF_ZZ_1; UNKNOWN_1.
DR PROSITE; PS01355; ZF_ZZ_2; 1.
SQ SEQUENCE 7 AA; 779 MW; 737728769DDDD6F0 CRC64;

Query Match 42.1%; Score 16; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.8e+06;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWX 3
DB 5 GLW 7

RESULT 4
EI01_LITRU STANDARD; PRT; 6 AA.
ID EI01_LITRU
AC P82056;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Electrin 1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog Litoria rubella.";
RL Aust. J. Chem. 52:639-645(1999).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Skin.
KW Amidation, Amphibian defense peptide; Direct protein sequencing.
FT MOD RES 6 792 MW; 6683704772C9A000 CRC64;
SQ SEQUENCE 6 AA; 792 MW; 6683704772C9A000 CRC64;

Query Match 39.5%; Score 15; DB 1; Length 6;
Best Local Similarity 20.0%; Pred. No. 1.8e+06;
Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXW 7
DB 1 FVPIW 5

RESULT 5
LOKI_LOCMI STANDARD; PRT; 6 AA.
ID LOKI_LOCMI
AC P41491;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Locustakinin I.
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=92262851; PubMed=1585017;
RA Schoofs L., Holman G.M., Proost P., van Damme J., Hayes T.K., de Loof A.;

RT "Locustakinin, a novel myotropic peptide from Locusta migratoria, isolation, primary structure and synthesis.";

RL Regul. Pept. 37:49-57(1992).

CC -1- FUNCTION: Myotropic peptide. May be important in the stimulation of ion transport and inhibition of diuretic activity in Malpighian tubules.

CC -1- SUBCELLULAR LOCATION: Secreted.

DR PIR; A61068; A61068.

KW Amidation; Direct protein sequencing; Neuropeptide.

FT MOD_RES 6 6 Glycine amide.

SQ SEQUENCE 6 AA; 654 MW; 686365A5B9CDB000 CRC64;

Query Match 36.8%; Score 14; DB 1; Length 6;

Best Local Similarity 25.0%; Pred. No. 1.8e+06;

Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 XXXX 7

Db : : : |

2 FSSW 5

RESULT 6

BRHP_CONIM STANDARD; PRT; 7 AA.

AC P58803;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Bromoheptapeptide Im.

OS Conus imperialis (Imperial cone).

OC Eukaryota; Metazoa; Gastropoda; Orthogastropoda;

OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;

OC Neogastropoda; Conoidea; Conidae; Conus.

OX NCBI_TaxID=35631;

RN [1]

RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.

RC TISSUE=Venom;

RX MEDLINE=97184108; PubMed=9030520;

RA Craig A.G., Jimenez E.C., Dykert J., Nielsen D.B., Gulyas J., Abogadie F.C., Porter J., Rivier J.E., Cruz L.J., Olivera B.M., McIntosh J.M.;

RT "A novel post-translational modification involving bromination of tryptophan. Identification of the residue, L-6-bromotryptophan, in peptides from Conus imperialis and Conus radiatus venom.";

RL J. Biol. Chem. 272:4689-4698(1997).

CC -1- FUNCTION: Does not elicit gross behavioral symptoms when injected centrally or peripherally in mice.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.

CC -1- MASS SPECTROMETRY: MW=853.19; METHOD=LSI; RANGE=1-7; NOTE=Ref.1.

DR PIR; A58512; A58512.

KW Amidation; Bromination; Direct protein sequencing;

KW Pyrrolidone carboxylic acid.

FT DISULFID 2 7 Pyrrolidone carboxylic acid.

FT MOD_RES 1 1 6'-bromotryptophan.

FT MOD_RES 6 6 Cysteine amide.

FT MOD_RES 7 7

SQ SEQUENCE 7 AA; 795 MW; 6EA3DC6D87EA6B0 CRC64;

Query Match 36.8%; Score 14; DB 1; Length 7;

Best Local Similarity 25.0%; Pred. No. 1.8e+06;

Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 XXXX 7

Db : : : |

3 GQAW 6

RESULT 7

TPFY_PACDA STANDARD; PRT; 7 AA.

ID TPFY_PACDA

AC P83455;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Tryptophyllin-1 (pdt-1).

OS Pachymedusa dactylosa (Giant mexican leaf frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;

OC Phyllomedusinae; Pachymedusa.

OX NCBI_TaxID=75988;

RN [1]

RP SEQUENCE, MASS SPECTROMETRY, HYDROXYLATION OF PRO-3, AND AMIDATION OF PRO-7.

RC TISSUE=Skin secretion;

RA Chen T.B., Orr D.F., Shaw C.;

RT "Pachymedusa dactylosa tryptophyllin-1 (pdt-1): structural characterization, pharmacological activity and cloning of precursor cDNA.";

RT Submitted (SEP-2002) to Swiss-Prot.

CC -1- FUNCTION: Myoactive. Has selective relaxing activity on vascular smooth muscle.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: Skin.

CC -1- MASS SPECTROMETRY: MW=809.2; METHOD=MALDI; RANGE=1-7; NOTE=Ref.1.

DR GO; GO:0005576; C:extracellular; NAS.

DR GO; GO:0045986; P:negative regulation of smooth muscle contractility; NAS.

KW Amidation; Amphibian defense peptide; Direct protein sequencing;

KW Hydroxylation.

FT MOD_RES 3 3 Hydroxyproline.

FT MOD_RES 7 7 Proline amide.

SQ SEQUENCE 7 AA; 794 MW; 7772D37DC776350 CRC64;

Query Match 36.8%; Score 14; DB 1; Length 7;

Best Local Similarity 25.0%; Pred. No. 1.8e+06;

Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 XXXX 7

Db : : : |

2 PPAW 5

RESULT 8

TVS1_LITRU STANDARD; PRT; 7 AA.

ID TVS1_LITRU

AC P82065;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Tryptophyllin 5.1.

OS Litoria rubella (Desert tree frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;

OC Pelodyadinae; Litoria.

OX NCBI_TaxID=104895;

RN [1]

RP SEQUENCE, AND MASS SPECTROMETRY.

RC TISSUE=Skin secretion;

RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C., Tyler M.J., Wallace J.C.;

RT "The structure of new peptides from the Australian red tree frog 'Litoria rubella'. The skin peptide profile as a probe for the study of evolutionary trends of amphibians.";

RT of evolutionary trends of amphibians.";

RL Aust. J. Chem. 49:955-963(1996).

CC -1- FUNCTION: May act as a neuromodulator or neurotransmitter.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.

CC -1- MASS SPECTROMETRY: MW=965; METHOD=FAB; RANGE=1-7; NOTE=Ref.1.

KW Amidation; Amphibian defense peptide; Direct protein sequencing;

KW Neuropeptide; Pyrrolidone carboxylic acid.

FT MOD_RES 1 1 Pyrrolidone carboxylic acid.

FT MOD_RES 7 7 Arginine amide.

SQ SEQUENCE 7 AA; 983 MW; 7401E9D3676046B0 CRC64;

Query Match 36.8%; Score 14; DB 1; Length 7;

Best Local Similarity 25.0%; Pred. No. 1.8e+06;

Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXX 6
|:::
Db 4 WFRH 7

RESULT 9

WWA1_ACHFU STANDARD; PRT; 7 AA.
ID WWA1_ACHFU
AC P35919;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Wwamide-1.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX NCBI_TaxID=6530;
RN [1]
RP SEQUENCE.
RC TISSUE=Ganglion; PubMed=8495720;
RX MEDLINE=93265912;
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT "Wwamide-1", -2 and -3: novel neuromodulatory peptides isolated from
RT ganglia of the African giant snail, Achatina fulica.";
RL FEBS Lett. 323:104-108(1993).
CC -1- FUNCTION: Exhibits modulatory effects on the peripheral nervous
CC system. Inhibits activity on a central neuron.
DR PIR; S33245; S33245.
DR PIR; S33245; S33245.
KW Amidation; Direct protein sequencing; Neuropeptide.
FT MOD_RES 7 7 Tryptophan amide.
SQ SEQUENCE 7 AA; 993 MW; 7362D5B69B041310 CRC64;

Query Match 36.8%; Score 14; DB 1; Length 7;
Best Local Similarity 25.0%; Pred. No. 1.8e+06;
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXX 6
|:::
Db 1 WREM 4

RESULT 10

WWA2_ACHFU STANDARD; PRT; 7 AA.
ID WWA2_ACHFU
AC P35920;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Wwamide-2.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX NCBI_TaxID=6530;
RN [1]
RP SEQUENCE.
RC TISSUE=Ganglion; PubMed=8495720;
RX MEDLINE=93265912;
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT "Wwamide-1", -2 and -3: novel neuromodulatory peptides isolated from
RT ganglia of the African giant snail, Achatina fulica.";
RL FEBS Lett. 323:104-108(1993).
DR PIR; S33246; S33246.
DR PIR; S33246; S33246.
KW Amidation; Direct protein sequencing; Neuropeptide.
FT MOD_RES 7 7 Tryptophan amide.
SQ SEQUENCE 7 AA; 964 MW; 7362D5B686D32310 CRC64;

Query Match 36.8%; Score 14; DB 1; Length 7;
Best Local Similarity 25.0%; Pred. No. 1.8e+06;
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXX 6
|:::
Db 1 WREM 4

Db |:::
1 WKQM 4

RESULT 11

WWA3_ACHFU STANDARD; PRT; 7 AA.
ID WWA3_ACHFU
AC P35921;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Wwamide-3.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX NCBI_TaxID=6530;
RN [1]
RP SEQUENCE.
RC TISSUE=Ganglion; PubMed=8495720;
RX MEDLINE=93265912;
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT "Wwamide-1", -2 and -3: novel neuromodulatory peptides isolated from
RT ganglia of the African giant snail, Achatina fulica.";
RL FEBS Lett. 323:104-108(1993).
DR PIR; S33244; S33244.
KW Amidation; Direct protein sequencing; Neuropeptide.
FT MOD_RES 7 7 Tryptophan amide.
SQ SEQUENCE 7 AA; 965 MW; 7362D5B69B132310 CRC64;

Query Match 36.8%; Score 14; DB 1; Length 7;
Best Local Similarity 25.0%; Pred. No. 1.8e+06;
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXX 6
|:::
Db 1 WKEM 4

RESULT 12

Q95945 PRELIMINARY; PRT; 7 AA.
ID Q95945
AC Q95945;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Inside intron 5 (Fragment).
OS Saccharomyces cerevisiae (Baker's yeast).
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D273-10B;
RX MEDLINE=81069885; PubMed=6254986;
RA Bonitz S.G., Coruzzi G., Thalenfeld B., Tzagoloff A., Macino G.;
RT "Assembly of the mitochondrial membrane system: Structure and
RT nucleotide sequence of the gene coding for subunit 1 of yeast
RT cytochrome oxidase.";
RL J. Biol. Chem. 255:11927-11941(1980).
DR EMBL; V00694; CAA24066.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 7 AA; 859 MW; 75B7232362CDC460 CRC64;

Query Match 36.8%; Score 14; DB 2; Length 7;
Best Local Similarity 25.0%; Pred. No. 1.8e+06;
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXX 6
|:::
Db 4 WKLS 7

RESULT 13
 O49223 PRELIMINARY; PRT; 7 AA.
 AC O49223;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HMg-1-like protein (Fragment).
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Roots;
 RX MEDLINE=91367679; PubMed=1891369;
 RA Laux T., Goldberg R.B.;
 RT "A plant DNA binding protein shares highly conserved sequence motifs
 with HMg-box proteins."
 RL Nucleic Acids Res. 19:4769-4769(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Roots;
 RA Mahalingam R., Knap H.T.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF047050; AAC03556.1; -.
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 850 MW; 6AAAAAB378637810 CRC64;

Query Match 36.8%; Score 14; DB 2; Length 7;
 Best Local Similarity 25.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
 Matches 1; Conservative 3; Mismatches 0;

QY 4 XXXW 7
 DB 1 GWGW 4

RESULT 14
 O8GL00 PRELIMINARY; PRT; 7 AA.
 AC O8GL00;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE PF-50 protein (Fragment).
 GN Name=PF-50;
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Plasmid group cp32-13.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CA15;
 RX MEDLINE=22610300; PubMed=12724373;
 RA Miller J.C., Stevenson B.;
 RT "Immunological and genetic characterization of Borrelia burgdorferi
 BapA and EppA proteins."
 RL Microbiology 149:1113-1125(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CA15;
 RX MEDLINE=22990544; PubMed=14629041;
 RA Stevenson B., Miller J.C.;
 RT "Intra- and interbacterial genetic exchange of Lyme disease spirochete
 erp genes generates sequence identity amidst diversity."
 RL J. Mol. Evol. 57:309-324(2003).
 DR EMBL; AY142106; AANI7857.1; -.
 KW Plasmid.
 FT NON_TER 1

SQ SEQUENCE 7 AA; 928 MW; 6337233050437350 CRC64;
 Query Match 36.8%; Score 14; DB 2; Length 7;
 Best Local Similarity 25.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
 Matches 1; Conservative 3; Mismatches 0;

QY 3 WXXX 6
 DB 2 WIIK 5

RESULT 15
 O8GL04 PRELIMINARY; PRT; 7 AA.
 AC O8GL04;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE PF-50 protein (Fragment).
 GN Name=PF-50;
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Plasmid group cp32-5.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=93-0107;
 RX MEDLINE=22610300; PubMed=12724373;
 RA Miller J.C., Stevenson B.;
 RT "Immunological and genetic characterization of Borrelia burgdorferi
 BapA and EppA proteins."
 RL Microbiology 149:1113-1125(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=93-0107;
 RX MEDLINE=22990544; PubMed=14629041;
 RA Stevenson B., Miller J.C.;
 RT "Intra- and interbacterial genetic exchange of Lyme disease spirochete
 erp genes generates sequence identity amidst diversity."
 RL J. Mol. Evol. 57:309-324(2003).
 DR EMBL; AY142103; AANI7848.1; -.
 KW Plasmid.
 FT NON_TER 1

SQ SEQUENCE 7 AA; 914 MW; 6337244330504310 CRC64;

Query Match 36.8%; Score 14; DB 2; Length 7;
 Best Local Similarity 25.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
 Matches 1; Conservative 3; Mismatches 0;

QY 3 WXXX 6
 DB 1 WIIK 4

Search completed: January 3, 2005, 17:26:03
 Job time : 187 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:11:20 ; Search time 90.88 Seconds
(without alignments)
31.578 Million cell updates/sec

Title: US-10-046-922-68

Perfect score: 39

Sequence: 1 GYVXXWX 8

Scoring table: BL6SUM62DX⁴

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: Genesecp1980s:*

2: Genesecp1990s:*

3: Genesecp2000s:*

4: Genesecp2001s:*

5: Genesecp2002s:*

6: Genesecp2003as:*

7: Genesecp2003bs:*

8: Genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	8	5	ABP53965 VEGFR-3 b
2	39	100.0	9	5	ABP53933 VEGFR-3 b
3	39	100.0	10	4	ABP53933 Rhesus D
4	39	100.0	10	5	ABP53931 VEGFR-3 b
5	39	100.0	10	5	ABP53968 VEGFR-3 b
6	39	100.0	10	5	ABP53932 VEGFR-3 b
7	39	100.0	11	2	AAW62148 Haemophil
8	39	100.0	11	8	ADJ25827 Tyrosine
9	39	100.0	12	4	ABP539769 Rhesus D
10	39	100.0	13	6	AAO26093 PC region
11	39	100.0	13	8	ADJ50760 Human ser
12	39	100.0	16	5	AAU93268 Granulocy
13	39	100.0	25	7	ADC99638 Cancer-re
14	39	100.0	29	4	ABB34417 Peptide #
15	39	100.0	29	4	AAW27898 Peptide #
16	39	100.0	29	4	ABB19829 Protein #
17	39	100.0	29	4	AAW55207 Human bra
18	39	100.0	35	4	ABB41036 Peptide #
19	39	100.0	35	4	AAW34810 Peptide #
20	39	100.0	35	4	ABB25111 Protein #
21	39	100.0	35	4	AAW74694 Human bon
22	39	100.0	35	4	AAW61892 Human bra
23	39	100.0	35	4	ABG56478 Human liv
24	39	100.0	35	5	ABG44505 Human pep
25	39	100.0	69	4	AAO13595 Human pol

26	39	100.0	74	4	AAU60558
27	39	100.0	74	6	ABM57077
28	39	100.0	91	4	AAU48276
29	39	100.0	91	6	ABM44795
30	39	100.0	100	3	ABM41381
31	39	100.0	116	2	AAK15437
32	39	100.0	116	6	ABO27261
33	39	100.0	116	6	ABO27269
34	39	100.0	116	6	ABO27263
35	39	100.0	116	6	ABO27259
36	39	100.0	116	6	ABO27255
37	39	100.0	116	6	ABO27277
38	39	100.0	116	6	ABO27273
39	39	100.0	116	6	ABO27257
40	39	100.0	116	6	ABO27271
41	39	100.0	116	6	ABO27267
42	39	100.0	117	6	ABO27265
43	39	100.0	118	3	AAW52191
44	39	100.0	118	7	ADJ95639
45	39	100.0	119	8	ADL11890

ALIGNMENTS

RESULT 1

ABP53965

ID ABP53965 standard; peptide; 8 AA.

AC ABP53965;

DT 09-JAN-2003 (first entry)

DE VEGFR-3 binding peptide SEQ ID NO:68.

KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
KW cytotropic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
KW vulnary; cell surface receptor; cancer; neovascularisation;
KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
KW diabetes; PDGF; platelet derived growth factor.

XX Homo sapiens.

OS Synthetic.

PH Key Location/Qualifiers

FT Misc-difference 4. .6 /note= "X is any amino acid"

FT Misc-difference 8 /note= "any amino acid"

WO200257299-A2.

25-JUL-2002.

16-JAN-2002; 2002WO-IB000099.

17-JAN-2001; 2001US-0262476P.

(LUDW-) LUDWIG INST CANCER RES.

(LICN) LICENTIA LTD.

PI Alitalo K, Koivunen E, Kubo H;

WPI; 2002-691521/74.

XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
PT such as cancer and diseases of neovascularization.

XX Claim 22; Page 81; 149pp; English.

CC The present invention describes an isolated peptide (I) that binds to and

CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnerary activities, and can be used in gene therapy.
 CC compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
 CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 XX Sequence 8 AA;
 SQ

Query Match 100.0%; Score 39; DB 5; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXWX 8
 Db 1 GYWXXXWX 8
 |||||

RESULT 2
 ABP53933
 ID ABP53933 standard; peptide; 9 AA.
 XX AC
 XX ABP53933;
 DT 09-JAN-2003 (first entry)
 XX VEGFR-3 binding peptide SEQ ID NO:36.
 XX Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
 KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
 KW cytosstatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
 KW vulnerary; cell surface receptor; cancer; neovascularisation;
 KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
 KW diabetes; PDGF; platelet derived growth factor.
 XX Homo sapiens.
 OS Synthetic.
 XX WO200257299-A2.
 PN 25-JUL-2002.
 XX 16-JAN-2002; 2002WO-IB000099.
 XX 17-JAN-2001; 2001US-0262476P.
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA (LICN) LICENTIA LTD.
 XX Alitalo K, Koivunen E, Kubo H;
 PI WPI; 2002-691521/74.
 XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
 PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
 PT such as cancer and diseases of neovascularization.
 XX Claim 14; Page 80; 149pp; English.
 XX The present invention describes an isolated peptide (I) that binds to and
 CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
 CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,
 CC antidiabetic and vulnerary activities, and can be used in gene therapy.
 CC compositions and methods from the present invention are useful for
 CC diagnosing, evaluating and treating disorders mediated by the activity of
 CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,
 CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
 CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
 CC chronic hepatitis, haemangiomas and diabetes. The present sequence
 CC represents a specifically claimed VEGFR-3 binding peptide from the
 CC present invention
 XX Sequence 9 AA;
 SQ

Query Match 100.0%; Score 39; DB 5; Length 9;
 Best Local Similarity 50.0%; Pred. No. 1.7e+06;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXWX 8
 Db 2 GYWDWTF 9
 |||:::

RESULT 3
 AAB99759
 ID AAB99759 standard; peptide; 10 AA.
 XX AC
 XX AAB99759;
 DT 21-SEP-2001 (first entry)
 XX Rhesus D antibody binding peptide SEQ ID NO:4.
 DE Rhesus D antibody binding peptide; Rhesus D; RHD; identification;
 KW anti-Rhesus D antibody; immunogen; epitope; diagnosis; therapy;
 KW prophylaxis; haemolytic disease of the newborn; HDN; ITP;
 KW idiopathic thrombocytopenic purpura; immunoglobulin.
 XX Homo sapiens.
 OS EP1106625-A1.
 PN 13-JUN-2001.
 XX 17-NOV-1999; 99EP-00122858.
 PF 17-NOV-1999; 99EP-00122858.
 PR 17-NOV-1999; 99EP-00122858.
 XX (ZLBB-) ZLB BIOTPLASMA AG.
 PA Miescher S, Hofmann A, Fisch I;
 PI WPI; 2001-383568/41.
 XX Novel peptides capable of binding Rhesus D antibodies are used to
 PT manufacture an agent for the diagnosis, therapy or prophylaxis of
 PT diseases associated with Rhesus D antigen, e.g. hemolytic disease of the
 PT newborn (HDN).
 XX Claim 1; Page 12; 19pp; English.
 XX The present sequence represents a peptide capable of binding Rhesus D
 CC antibodies (I). Also described in the present invention are: (1) a
 CC nucleic acid (II) encoding (I); (2) a vector (III) comprising one or more
 CC (II) operably linked to an expression control system; (3) a cell (IV)
 CC comprising (II) or (III); (4) preparing (I); (5) identifying (MI)
 CC peptides having immunologic properties of Rhesus D protein epitopes
 CC comprising subjecting an antibody/antibody fragment recognising an
 CC epitope of Rhesus D protein to several panning rounds with a phage
 CC display library, and identifying immunogenic peptide sequences which are
 CC mimotopes which differ in their amino acid sequence from the amino acid
 CC sequences of Rhesus D protein; and (6) peptides (V) with immunological
 CC properties of Rhesus D protein epitopes obtained by (MI). (I) is used to
 CC manufacture an agent for the diagnosis, therapy or prophylaxis of
 CC diseases associated with Rhesus D antigen, e.g. haemolytic disease of the
 CC newborn (HDN) or idiopathic thrombocytopenic purpura (ITP), for the
 CC manufacture of an affinity reagent for anti-Rhesus D antibodies purified

CC or removed from body fluids or immunoglobulin preparations. Using (I) as
CC an immunogen to raise anti-rhesus D antibodies avoids using immunisation
CC with foreign erythrocytes thereby avoiding the risk of transmission of
CC viral diseases like AIDS and hepatitis B

XX SQ Sequence 10 AA;
Query Match 100.0%; Score 39; DB 4; Length 10;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXX 8
|||::|:
Db 1 GYWSAKWA 8

RESULT 4
ABP53931
ID ABP53931 standard; peptide; 10 AA.

XX AC ABP53931;

DT 09-JAN-2003 (first entry)

DE VEGFR-3 binding peptide SEQ ID NO:34.

XX KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
KW cytosolic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
KW vulnary; cell surface receptor; cancer; neovascularisation;
KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
KW diabetes; PDGF; platelet derived growth factor.

XX OS Homo sapiens.
OS Synthetic.

XX FH Key Location/Qualifiers
FT Misc-difference 1 /note= "any amino acid"
FT Misc-difference 10 /note= "any amino acid"
FT Misc-difference 10 /note= "any amino acid"

XX PN WO200257299-A2.

XX PD 25-JUL-2002.

XX PF 16-JAN-2002; 2002WO-IB000099.

XX PR 17-JAN-2001; 2001US-0262476P.

XX PA (LUDW-) LUDWIG INST CANCER RES.
XX PA (LICN) LICENTIA LTD.

XX PI Alitalo K, Koivunen E, Kubo H;

XX WPI; 2002-691521/74.

XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
PT such as cancer and diseases of neovascularization.

XX Claim 12; Page 80; 149pp; English.

XX CC The present invention describes an isolated peptide (I) that binds to and
CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
CC have cytosolic, hepatotropic, antiinflammatory, hypotensive,
CC antidiabetic and vulnerary activities, and can be used in gene therapy.
CC Compositions and methods from the present invention are useful for
CC diagnosing, evaluating and treating disorders mediated by the activity of
CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
CC liver, spleen, kidney, lymph node, small intestine, blood cells,
CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of

CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
CC chronic hepatitis, haemangiomas and diabetes. The present sequence
CC represents a specifically claimed VEGFR-3 binding peptide from the
CC present invention

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 39; DB 5; Length 10;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXX 8
|||::|:
Db 2 GYWLTIWG 9

RESULT 5
ABP53968
ID ABP53968 standard; peptide; 10 AA.

XX AC ABP53968;

DT 09-JAN-2003 (first entry)

DE VEGFR-3 binding peptide SEQ ID NO:73.

XX KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
KW cytosolic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
KW vulnary; cell surface receptor; cancer; neovascularisation;
KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
KW diabetes; PDGF; platelet derived growth factor.

XX OS Homo sapiens.
OS Synthetic.

XX FH Key Location/Qualifiers
FT Misc-difference 5.7 /note= "X is any amino acid"
FT Misc-difference 9 /note= "X is any amino acid"
FT Misc-difference 9 /note= "X is any amino acid"

XX PN WO200257299-A2.

XX PD 25-JUL-2002.

XX PF 16-JAN-2002; 2002WO-IB000099.

XX PR 17-JAN-2001; 2001US-0262476P.

XX PA (LUDW-) LUDWIG INST CANCER RES.
XX PA (LICN) LICENTIA LTD.

XX PI Alitalo K, Koivunen E, Kubo H;

XX WPI; 2002-691521/74.

XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
PT such as cancer and diseases of neovascularization.

XX Disclosure; Page 147; 149pp; English.

XX CC The present invention describes an isolated peptide (I) that binds to and
CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
CC have cytosolic, hepatotropic, antiinflammatory, hypotensive,
CC antidiabetic and vulnerary activities, and can be used in gene therapy.
CC Compositions and methods from the present invention are useful for
CC diagnosing, evaluating and treating disorders mediated by the activity of
CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
CC liver, spleen, kidney, lymph node, small intestine, blood cells,
CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of

CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,
CC chronic hepatitis, haemangiomas and diabetes. The present sequence
CC represents a VEGFR-3 binding peptide, which is given in the
CC exemplification of the present invention
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 39; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXWX 8
Db 2 GYXXXXWX 9
|||||

RESULT 6
ABP53932
ID ABP53932 standard; peptide; 10 AA.
XX
AC ABP53932;
XX
DT 09-JAN-2003 (first entry)
XX
DE VEGFR-3 binding peptide SEQ ID NO:35.
XX
KW Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3;
KW angiogenesis; lymphangiogenesis; vascular endothelial growth factor;
KW cytostatic; hepatotropic; antiinflammatory; hypotensive; antidiabetic;
KW vulnerary; cell surface receptor; cancer; neovascularisation;
KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma;
KW diabetes; PDGF; platelet derived growth factor.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200257299-A2.
XX
PD 25-JUL-2002.
XX
PF 16-JAN-2002; 2002WO-1B000099.
XX
PR 17-JAN-2001; 2001US-0262476P.
XX
PA (JUDW-) LUDWIG INST CANCER RES.
PA (LICN) LICENTIA LTD.
XX
PI Alitalo K, Koivunen E, Kubo H;
XX
DR WPI; 2002-691521/74.
XX
PT New isolated peptide that inhibits VEGF-C and VEGF-D, useful for
PT diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity,
PT such as cancer and diseases of neovascularization.
XX
PS Claim 13; Page 80; 149pp; English.
XX

The present invention describes an isolated peptide (I) that binds to and
inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)
have cytostatic, hepatotropic, antiinflammatory, hypotensive,
antiadipetic and vulnerary activities, and can be used in gene therapy.
Compositions and methods from the present invention are useful for
diagnosing, evaluating and treating disorders mediated by the activity of
the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,
liver, spleen, kidney, lymph node, small intestine, blood cells,
pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,
skin, head and neck, oesophagus, bone, marrow or blood, and diseases of
neovascularisation, e.g. liver diseases, hypertension, post-trauma,
chronic hepatitis, haemangiomas and diabetes. The present sequence
represents a specifically claimed VEGFR-3 binding peptide from the
present invention
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 39; DB 5; Length 10;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXWX 8
Db 2 GYWLTIWG 9
|||||

RESULT 7
AAW62148
ID AAW62148 standard; peptide; 11 AA.
XX
AC AAW62148;
XX
DT 16-SEP-1998 (first entry)
XX
DE Haemophilus influenzae tyrosine tRNA synthetase binding peptide 21.
XX
KW Identification; ligand; biological activity; target-binding;
KW drug screening; library; inhibitory ligand.
XX
OS Synthetic.
OS Haemophilus influenzae.
XX
PN WO9819162-A1.
XX
PD 07-MAY-1998.
XX
PF 31-OCT-1997; 97WO-US019638.
XX
PR 31-OCT-1996; 96US-00740671.
XX
PA (NOVA-) NOVALON PHARM CORP.
XX
PI Fowlkes DM, Kay BK, Frelinger JA, Hyde-Deruysscher RP;
XX
DR WPI; 1998-272389/24.
XX

Identifying ligands which mediate biological activity of a protein - by
identifying target-binding ligands and screening a library for ligands
which inhibit target-binding ligand mediated activity.

Example 5; Page 100; 143pp; English.

A method has been developed for identifying a ligand which mediates the
biological activity of a target protein (T) by inhibiting the binding of
(T) to a binding partner. The method comprises: (a) screening a first
combinatorial library comprising first member ligands for binding to the
target-binding ligands (TBLs), to identifying one or more TBLs; (b)
screening a second library comprising second member ligands for the
ability to inhibit the binding of one or more of the TBLs to the target
protein, and so obtaining one or more inhibitory ligands; and (c)
determining which of the inhibitory ligands can mediate a biological
activity of the target protein. The present sequence represents a
potential binding peptide for Haemophilus influenzae tyrosine tRNA
synthetase from an example of the present invention. The method can be
used for identifying drugs which can mediate the biological activity of a
target protein. It can be used to identify the biological activity of a
target protein whose biological function is not known and perhaps cannot
be determined directly. The method can also be used to identify new
inhibitory ligands of specific target proteins. The method provides high
throughput screens which are essentially identical for similar and
dissimilar targets, bypassing the need to develop distinct assays for
biochemically diverse targets

Sequence 11 AA;

Query Match 100.0%; Score 39; DB 2; Length 11;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8
 DB 1 GYWPDPWG 8

RESULT 8

ADJ25827
 ID ADJ25827 standard; peptide; 11 AA.

AC ADJ25827;
 XX

DT 20-MAY-2004 (first entry)

DE Tyrosine tRNA synthetase binding peptide #21.

XX ligand identification; peptide library;

KW complementary combinatorial library; tyrosine tRNA synthetase.

XX Synthetic.

PN US6617114-B1.

PD 09-SEP-2003.

PF 30-APR-1998; 98US-00069827.

PR 31-OCT-1996; 96US-00740671.

PR 31-OCT-1997; 97WO-US019638.

PR 31-MAR-1998; 98US-00050359.

XX (KARO-) KARO BIO AB.

PI Fowlkes DM, Kay BK, Frelinger JA, Hyde-Deruyser RP;

XX WPI; 2004-068186/07.

XX Identification of ligand that can mediate biological activity of target
 PT protein, comprises screening first combinatorial library having first
 PT member ligands for binding to target protein to identify target-binding
 PT ligand(s).

XX Example 5; SEQ ID NO 87; 98pp; English.

XX The invention relates to a method of identifying a ligand that can
 CC mediate the biological activity of target protein via inhibition of the
 CC binding of target protein to a binding partner ligand comprising
 CC screening first combinatorial library having first member ligands for
 CC binding to target protein to identify target-binding ligand(s). The
 CC method is useful for identifying ligands that can mediate the biological
 CC activity of target proteins via inhibition of the binding of target
 CC protein to a binding partner ligand. The invention does not require that
 CC the natural binding partner be used as reagent. The need for the natural
 CC binding partner is obviated with the use of complementary combinatorial
 CC libraries. The present sequence is used in the exemplification of the
 CC present invention.

XX Sequence 11 AA;

Query Match 100.0%; Score 39; DB 8; Length 11;
 Best Local Similarity 50.0%; Pred. No. 51;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8
 DB 1 GYWPDPWG 8

RESULT 9

AAB99769
 ID AAB99769 standard; peptide; 12 AA.

XX AAB99769;

XX

DT 21-SEP-2001 (first entry)

XX Rhesus D antibody related peptide #5.

XX Rhesus D antibody binding peptide; Rhesus D; RHD; identification;
 KW anti-Rhesus D antibody; immunogen; epitope; diagnosis; therapy;
 KW prophylaxis; haemolytic disease of the newborn; HDN; ITP; cyclic;
 KW idiopathic thrombocytopenic purpura; immunoglobulin; circular.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 1..12

PN EP1106625-A1.

XX 13-JUN-2001.

XX 17-NOV-1999; 99EP-00122858.

XX 17-NOV-1999; 99EP-00122858.

XX (ZLBB-) ZLB BIOPLASMA AG.

XX Miescher S, Hofmann A, Fisch I;

XX WPI; 2001-383568/41.

XX Novel peptides capable of binding Rhesus D antibodies are used to
 PT manufacture an agent for the diagnosis, therapy or prophylaxis of
 PT diseases associated with Rhesus D antigen, e.g. hemolytic disease of the
 PT newborn (HDN).

XX Example 1; Page 8; 19pp; English.

XX The present invention describes peptides capable of binding Rhesus D
 CC antibodies (I). Also described in the present invention are: (1) a
 CC nucleic acid (II) encoding (I); (2) a vector (III) comprising one or more
 CC (II) operably linked to an expression control system; (3) a cell (IV)
 CC comprising (II) or (III); (4) preparing (I); (5) identifying (MI)
 CC peptides having immunologic properties of Rhesus D protein epitopes
 CC comprising subjecting an antibody/antibody fragment recognising an
 CC epitope of Rhesus D protein to several immunogenic peptide sequences which are
 CC display library, and identifying immunogenic peptide sequences which are
 CC mimotopes which differ in their amino acid sequence from the amino acid
 CC sequences of Rhesus D protein; and (6) peptides (V) with immunological
 CC properties of Rhesus D protein epitopes obtained by (MI). (I) is used to
 CC manufacture an agent for the diagnosis, therapy or prophylaxis of
 CC diseases associated with Rhesus D antigen, e.g. haemolytic disease of the
 CC newborn (HDN) or idiopathic thrombocytopenic purpura (ITP), for the
 CC manufacture of an affinity reagent for anti-Rhesus D antibodies purified
 CC or removed from body fluids or immunoglobulin preparations. Using (I) as
 CC an immunogen to raise anti-Rhesus D antibodies avoids using immunisation
 CC with foreign erythrocytes thereby avoiding the risk of transmission of
 CC viral diseases like AIDS and hepatitis B. The present sequence represents
 CC an anti-Rhesus D (RHD) antibody related peptide which is used in an
 CC example from the present invention

XX Sequence 12 AA;

Query Match 100.0%; Score 39; DB 4; Length 12;
 Best Local Similarity 50.0%; Pred. No. 56;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8
 DB 2 GYWSAKWA 9

RESULT 10

AAO26093

ID AAO26093 standard; peptide; 13 AA.

XX AAO26093;
AC
XX
DT 03-APR-2003 (first entry)
XX
DE Fe region binding peptide SEQ ID No 73.
XX
KW Immunoglobulin Fc region; binding; whole blood; plasma; transgenic milk;
KW antibody response; half-life; stability; circulatory system.
XX
OS Unidentified.
XX
PN WO200286070-A2.
XX
PD 31-OCT-2002.
XX
PF 18-APR-2002; 2002WO-US012492.
XX
PR 18-APR-2001; 2001US-0284534P.
XX
PA (DYAX-) DYAX CORP.
XX
PI Rondon IJ, Wu Q, Ley AC, Stochl M, Ransohoff TC, Potter MD;
XX WPI; 2003-201220/19.
XX
XX New polypeptides, useful as binding molecules for detecting, isolating or
PT purifying immunoglobulin Fc-region polypeptides present in a solution, or
PT for regulating or preventing an antibody response.
XX
PS Claim 3; Page 76; 152pp; English.
XX
CC The invention relates to novel isolated polypeptides comprising a
CC sequence that binds an immunoglobulin Fc region. The polypeptides are
CC useful as binding molecules for detecting, isolating or purifying
CC immunoglobulin Fc-region polypeptides present in a solution, e.g. whole
CC blood, plasma or transgenic milk. The Fc-region binding polypeptides are
CC also useful for regulating or preventing an antibody response, or for
CC increasing the half-life and over all stability of a therapeutic or
CC diagnostic compound that is administered to or enters the circulatory
CC system of an individual. This sequence represents an Fc region binding
CC peptide of the invention
XX
XX Sequence 13 AA;
XX
Query Match 100.0%; Score 39; DB 6; Length 13;
Best Local Similarity 50.0%; Pred. No. 60;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYVXXWX 8
DB |||::|:
1 GYWCNVWG 8
RESULT 11
ADJ50760
ID ADJ50760 standard; peptide; 13 AA.
XX
AC ADJ50760;
XX
XX 06-MAY-2004 (first entry)
DT
XX Human serum albumin binding peptide, Seq ID No 297.
DE
XX human serum albumin; HSA; serum; blood; tumour; human.
KW
XX Homo sapiens.
OS
XX WO2003106493-A1.
PN
XX
PD 24-DEC-2003.
XX
XX 16-JUN-2003; 2003WO-US018896.
PF

XX 14-JUN-2002; 2002US-0388642P.
PR (DYAX-) DYAX CORP.
XX
PI Sato AK, Dawson BM;
XX
DR WPI; 2004-082161/08.
XX
PT Evaluating sample comprising soluble serum protein by forming complex
PT comprising serum protein and physically associated compounds using
PT peptide ligand that specifically binds with proteins, which is separated
PT and evaluated.
XX
PS Disclosure; SEQ ID NO 297; 191pp; English.
XX
CC The invention relates to a method of evaluating sample by providing a
CC soluble serum protein (I), one or more compounds physically associated
CC with (I), and a (I)-binding agent that comprises a peptide that
CC specifically binds to (I), allowing the (I)-binding agent to bind to (I)
CC to form a complex including one or more compounds physically associated
CC with (I), separating the complex from one or more components of the
CC sample, and evaluating one or more of the physically associated
CC compounds. The sample comprises blood or serum, or is obtained from a
CC biopsy. The sample may also be obtained from a tumour or a region within
CC 5 mm of a tumour. The method is useful for detecting modulators that
CC modulate interaction of serum protein-binding compound and serum protein
CC and for identifying binding ligands for serum protein. The present
CC sequence represents a serum albumin-binding peptide identified using the
CC method of the invention.
XX
XX Sequence 13 AA;
XX
Query Match 100.0%; Score 39; DB 8; Length 13;
Best Local Similarity 50.0%; Pred. No. 60;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYVXXWX 8
DB |||::|:
1 GYWCNVWG 8
RESULT 12
AAU93268
ID AAU93268 standard; peptide; 16 AA.
XX
AC AAU93268;
XX
DT 02-JUL-2002 (first entry)
XX
DE Granulocyte-colony stimulating factor receptor binding peptide #74.
XX
KW G-CSFR; granulocyte-colony stimulating factor receptor; cytokine;
KW haematopoietic growth factor; neutrophil proliferation; AIDS;
KW neutrophil differentiation; acquired immunodeficiency syndrome;
KW chemotherapy-induced neutropaenia; community acquired pneumonia;
KW depressed neutrophil count; immunostimulant.
XX
OS Synthetic.
XX
PN WO200207676-A2.
XX
PD 31-JAN-2002.
XX
PF 20-JUL-2001; 2001WO-US023046.
XX
PR 20-JUL-2000; 2000US-00620091.
XX
PA (GLAXO) GLAXO GROUP LTD.
XX
PI Cwirila SE, Balu P, Duffin DJ, Piplani S, Mcsowen-Merrill B;
PI Schatz PJ;
XX

DR WPI; 2002-329382/36.
XX Novel compounds, useful for treating depressed neutrophil count, comprise
PT peptide chains of approximately 6 to 40 amino acids in length that bind
PT to granulocyte-colony stimulating factor receptor.
XX
PS Claim 4; Page 53; 90pp; English.
XX
XX The invention relates to compounds comprising a peptide chain
CC approximately 6 to 40 amino acids in length that binds to granulocyte-
CC colony stimulating factor receptor (G-CSFR). The compounds contain
CC specific sequences of the generic peptides appearing as AAU79402-AAU79406
CC and the generic sequences XV 1XV 2XV 3XV 4XV 5XV 6XV 7XV 8 (where XV 1 =
CC E, C, Q, V or Y; XV 2 = E, A, L, M, S, W or Q; XV 3 = K, R or T; XV 4 =
CC L, A or V; XV 5 = R, A, M, H, E, V, L, G, D, O or S; XV 6 = E or V; XV 7
CC = A or G; and XV 8 = R, H, G or L) and XVI 1XVI 2XVI 3XVI 4XVI 5
CC EXVI 6XVI 7XVI 8XVI 9 (where XVI 1 = A, E or G; XVI 2 = E, H or D; XVI 3
CC = R or G; XVI 4 = K, Y, M, N, Q, R, D, I, S or E; XVI 5 = A, S or P;
CC XVI 6 = E, D, T, Q, K or A; XVI 7 = R, W, K, L, S, A or Q; XVI 8 = R or E
CC ; and XVI 9 = W, G or R). The compounds are used for treating conditions
CC associated with depressed neutrophil count e.g. chemotherapy induced
CC neutropenia, AIDS-induced neutropenia or community-acquired pneumonia-
CC induced pneumonia. The compounds are useful as in vitro as tools for
CC understanding the biological role of granulocyte-colony stimulating
CC factor (G-CSF) a haematopoietic growth factor and cytokine that stimulates
CC neutrophil proliferation and differentiation), including evaluation of
CC many factors thought to influence, and be influenced by, production of
CC white blood cells, in the development of compounds that bind to G-CSFR,
CC as reagents for detecting G-CSF receptor or related receptor on living
CC cells, fixed cells, in biological fluid, in tissue homogenates or in
CC purified natural biological materials, in situ staining, fluorescence-
CC activated cell sorting (FACS), Western blotting or enzyme-linked
CC immunoadsorbent assay (ELISA), in receptor purification or in purifying
CC cells expressing G-CSFR on the cell surface (or inside permeabilised
CC cells) as a commercial research reagent for various medical and
CC diagnostic uses or to treat a disease that would benefit from the ability
CC to of a compound to mimic the effects of G-CSF in vivo. The compounds
CC bind specifically to G-CSFR and allow for studies of biological
CC activities mediated by the receptor and for the treatment of diseases,
CC disorders and conditions that would benefit from activating or
CC inactivating G-CSFR. The present sequence is a G-CSFR binding peptide of
CC the invention
XX
SQ Sequence 16 AA;
Query Match 100.0%; Score 39; DB 5; Length 16;
Best Local Similarity 50.0%; Pred. No. 74;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYVXXWX 8
Db |||:::|
1 GYVCDPWT 8
RESULT 13
ADC99638
ID ADC99638 standard; peptide; 25 AA.
XX
AC ADC99638;
XX
DT 01-JAN-2004 (first entry)
XX
DE Cancer-related Tie-1-binder peptide - SEQ ID 476.
XX
KW cytostatic; cancer; gene therapy; DGI-2; DGI-5; DGI-7; DGI-9; Hras;
KW leptin; VEGF; vascular endothelial growth factor receptor; VEGF-R1;
KW VEGF-R2; VEGF-R3; FLTL; FMS-related tyrosine kinase 1; FLK1; KDR;
KW kinase insert domain protein receptor; EGFR; epidermal growth factor;
KW FGFR1; fibroblast growth factor; Tie-1.
XX
OS Unidentified.
XX
PN WO2003035839-A2.

XX 01-MAY-2003.
PD
XX 24-OCT-2002; 2002WO-US034021.
PF
XX 24-OCT-2001; 2001US-0345471P.
PR
XX (DGI-B-) DGI BIOTECHNOLOGIES INC.
PA
XX Pillutla RC, Brissette R, Spruyt M, Dedova O, Blume A;
PI Prendergast J, Goldstein N;
DR WPI; 2003-457332/43.
XX
XX Selecting target and target binder pairs for preparing a composition for
PT treating cancer by mixing in a reaction vessel phage expressing
PT biological targets and phage expressing target binders.
XX
PS Claim 26; SEQ ID NO 476; 172pp; English.
XX
CC The invention relates to a novel method of selecting target and target
CC binder pairs comprising mixing in a reaction vessel phage expressing
CC biological targets and phage expressing target binders, each having
CC distinguishable selection markers and selecting target and target binder
CC pairs based on the selection markers. The molecules of the invention
CC demonstrate cytostatic activity whilst the method may be useful for
CC selecting target and target binder pairs for preparing a composition for
CC treating cancer. Furthermore, the method may be utilised during gene
CC therapy procedures. The current sequence is that of the cancer-related
CC Tie-1-binder peptide of the invention.
XX
SQ Sequence 25 AA;
Query Match 100.0%; Score 39; DB 7; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYVXXWX 8
Db |||:::|
6 GYVGLWNG 13
RESULT 14
ABB34417
ID ABB34417 standard; peptide; 29 AA.
XX
AC ABB34417;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #1923 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI

XX WPI; 2001-483447/52.
XX
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
XX Claim 27; SEQ ID NO 27052; 639pp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 29 AA;
SQ

Query Match 100.0%; Score 39; DB 4; Length 29;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYWXWXX 8
Db 17 GYWLHWM 24

RESULT 15
AAM27898
ID AAM27898 standard; protein; 29 AA.
XX
XX AAM27898;
XX
XX 17-OCT-2001 (first entry)
XX
XX Peptide #1935 encoded by probe for measuring placental gene expression.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-48897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
XX Claim 27; SEQ ID NO 28167; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
CC see AA131315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:19:47 ; Search time 23.04 Seconds
(without alignments)
23.027 Million cell updates/sec

Title: US-10-046-922-68

Perfect score: 39

Sequence: 1 GYVXXWX 8

Scoring table: -BLOSUM62DX²

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*

2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*

3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*

4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*

5: /cgn2_6/prodata/1/1aa/PCUS_COMB.pep:*

6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	11	4	US-09-069-827A-87
2	39	100.0	16	4	US-09-620-091-81
3	39	100.0	177	4	US-09-543-681A-7620
4	39	100.0	252	3	US-09-502-653-10
5	39	100.0	263	4	US-09-610-906-12
6	39	100.0	274	4	US-09-248-796A-15791
7	39	100.0	278	4	US-09-145-828A-11
8	39	100.0	278	4	US-09-903-456-18
9	39	100.0	339	4	US-09-252-991A-26841
10	39	100.0	342	4	US-09-252-991A-32031
11	39	100.0	362	1	US-08-415-751-6
12	39	100.0	367	4	US-09-248-796A-15188
13	39	100.0	478	4	US-09-107-532A-8668
14	39	100.0	492	4	US-09-107-532A-6945
15	39	100.0	499	4	US-09-252-991A-23328
16	39	100.0	500	4	US-09-252-991A-21214
17	39	100.0	543	4	US-09-252-991A-18697
18	39	100.0	865	4	US-09-252-991A-19339
19	39	100.0	1498	4	US-09-792-616-9
20	39	100.0	1503	4	US-09-792-616-3
21	35	89.7	34	4	US-09-270-767-60715
22	35	89.7	43	2	US-08-488-161-69
23	35	89.7	43	3	US-09-273-685-69
24	35	89.7	43	5	PCR-US95-11934-69
25	35	89.7	55	4	US-09-621-976-7633
26	35	89.7	101	3	US-09-374-135-4
27	35	89.7	158	4	US-09-270-767-35277

28 35 89.7 158 4 US-09-270-767-50494 Sequence 50494, A

29 35 89.7 197 3 US-09-112-248-2 Sequence 2, Appli

30 35 89.7 227 4 US-09-489-039A-8752 Sequence 8752, Ap

31 35 89.7 233 4 US-09-248-796A-20388 Sequence 20388, A

32 35 89.7 255 4 US-09-270-767-45223 Sequence 45223, A

33 35 89.7 328 4 US-09-452-937A-34 Sequence 34, Appli

34 35 89.7 417 4 US-09-248-796A-18939 Sequence 18939, A

35 35 89.7 458 4 US-09-543-681A-6324 Sequence 6324, Ap

36 35 89.7 519 4 US-09-198-452A-479 Sequence 479, App

37 35 89.7 580 4 US-09-270-767-43086 Sequence 43086, A

38 35 89.7 602 4 US-09-248-796A-16589 Sequence 16589, A

39 35 89.7 625 4 US-09-252-991A-28537 Sequence 28537, A

40 35 89.7 639 4 US-09-270-767-61294 Sequence 61294, A

41 35 89.7 764 4 US-09-270-767-61294 Sequence 45772, A

42 34 87.2 13 4 US-09-069-827A-86 Sequence 86, Appli

43 34 87.2 19 4 US-09-794-5298-8 Sequence 8, Appli

44 34 87.2 19 4 US-09-794-517A-8 Sequence 8, Appli

45 34 87.2 19 4 US-09-011-645E-8 Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-09-069-827A-87
; Sequence 87, Application US/09069827A
; Patent No. 6617114
; GENERAL INFORMATION:
; APPLICANT: FOWLKES, Dana M
; KAY, Brian K
; FRELINGER, Jeffrey A
; HYDE-DERUYSCHE, Robin P
; TITLE OF INVENTION: IDENTIFICATION OF DRUGS USING
; COMPLEMENTARY COMBINATORIAL LIBRARIES
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/069,827A
; FILING DATE: 30-Apr-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/050,359
; FILING DATE: 31-MAR-1998
; APPLICATION NUMBER: PCT/US97/19638
; FILING DATE: 31-OCT-1997
; APPLICATION NUMBER: US 08/740,671
; FILING DATE: 31-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Iver P
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: FOWLKES=4C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 87:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 87:

US-09-069-827A-87

Query Match 100.0%; Score 39; DB 4; Length 11;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXX 8
DB 1 GYWPDWG 8

RESULT 2

US-09-620-091-81
; Sequence 81, Application US/09620091
; Patent No. 6716811
; GENERAL INFORMATION:
; APPLICANT: CWIRLA, STEVEN E.
; APPLICANT: BALU, PALANI
; APPLICANT: DUFFIN, DAVID J.
; APPLICANT: PIPLANI, SUNILA
; APPLICANT: MERRILL, BARBARA MCEOWEN
; APPLICANT: SCHATZ, PETER JOSEPH
; TITLE OF INVENTION: COMPOUNDS HAVING AFFINITY FOR THE GRANULOCYTE-COLONY
; TITLE OF INVENTION: STIMULATING FACTOR RECEPTOR (G-CSFR) AND ASSOCIATED
; FILE REFERENCE: 0300-0014
; CURRENT APPLICATION NUMBER: US/09/620,091
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 491
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-620-091-81

Query Match 100.0%; Score 39; DB 4; Length 16;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXX 8
DB 1 GYWCDPWT 8

RESULT 3

US-09-543-681A-7620
; Sequence 7620, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7620
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7620

Query Match 100.0%; Score 39; DB 4; Length 177;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXX 8

DB 43 GYWINWQ 50

RESULT 4

US-09-502-653-10
; Sequence 10, Application US/09502653
; Patent No. 6331426
; GENERAL INFORMATION:
; APPLICANT: Bj rnvad, Mads Eskelund
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Schleim, Martin
; APPLICANT: Bech, Lisbeth
; APPLICANT: stergaard, Peter Rahbek
; APPLICANT: Sj holm, Carsten
; TITLE OF INVENTION: NOVEL GALACTANASES
; FILE REFERENCE: 5481.200-US
; CURRENT APPLICATION NUMBER: US/09/502,653
; CURRENT FILING DATE: 2000-02-11
; EARLIER APPLICATION NUMBER: PA 1999 00184
; EARLIER FILING DATE: 1999-02-11
; EARLIER APPLICATION NUMBER: PA 1999 00799
; EARLIER FILING DATE: 1999-06-07
; EARLIER APPLICATION NUMBER: 60/125,885
; EARLIER FILING DATE: 1999-03-24
; EARLIER APPLICATION NUMBER: 60/138,445
; EARLIER FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Bacillus circulans
US-09-502-653-10

Query Match 100.0%; Score 39; DB 3; Length 252;
Best Local Similarity 50.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXX 8
DB 210 GYWEPAWI 217

RESULT 5

US-09-610-906-12
; Sequence 12, Application US/09610906
; Patent No. 6566066
; GENERAL INFORMATION:
; APPLICANT: Walker, Michael G.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Klinger, Tod M.
; TITLE OF INVENTION: AQUAPORIN-8 VARIANT
; FILE REFERENCE: PC-0012 CIP
; CURRENT APPLICATION NUMBER: US/09/610,906
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 09/226,994
; PRIOR FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. 6566066: g2346968
; PUBLIC INFORMATION:
US-09-610-906-12

Query Match 100.0%; Score 39; DB 4; Length 263;
Best Local Similarity 50.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
; Patent No. 6677145
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407 US P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-09-903-456-18

Query Match      100.0%; Score 39; DB 4; Length 278;
Best Local Similarity 50.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 GYXXXXXX 8
DB      108 GYWIFLMV 115

RESULT 9
US-09-252-991A-26841
; Sequence 26841, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26841
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-26841

Query Match      100.0%; Score 39; DB 4; Length 339;
Best Local Similarity 50.0%; Pred. No. 5.8e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 GYXXXXXX 8
DB      261 GYGGGFWG 268

RESULT 10
US-09-252-991A-32031
; Sequence 32031, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
```

```
; Patent No. 6677145
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407 US P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-09-903-456-18

Query Match      100.0%; Score 39; DB 4; Length 278;
Best Local Similarity 50.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 GYXXXXXX 8
DB      108 GYWIFLMV 115

RESULT 6
US-09-248-796A-15791
; Sequence 15791, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15791
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-15791

Query Match      100.0%; Score 39; DB 4; Length 274;
Best Local Similarity 50.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 GYXXXXXX 8
DB      84 GYWPITWK 91

RESULT 7
US-09-145-828A-11
; Sequence 11, Application US/09145828A
; Patent No. 6403349
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Thurmond, Jennifer
; APPLICANT: Kirchner, Stephen J.
; APPLICANT: Parker-Barnes, Jennifer M.
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
; FILE REFERENCE: 6407 US O1
; CURRENT APPLICATION NUMBER: US/09/145,828A
; CURRENT FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-09-145-828A-11

Query Match      100.0%; Score 39; DB 4; Length 278;
Best Local Similarity 50.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 GYXXXXXX 8
DB      108 GYWIFLMV 115

RESULT 8
US-09-903-456-18
; Sequence 18, Application US/09903456
```

```
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32031
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32031

Query Match      100.0%; Score 39; DB 4; Length 342;
Best Local Similarity 50.0%; Pred. No. 5.9e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 GYWXWXXW 8
DB      315 GYWRGPMR 322

RESULT 11
US-08-415-751-6
; Sequence 6, Application US/08415751
; Patent No. 5643772
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
; TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA
; TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
; TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
; TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS AND KIT
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PHILLIPS, MOORE, LEMPION & FINLEY
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: California
; COUNTRY: United States of America
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
; COMPUTER: PC
; OPERATING SYSTEM: DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/415,751
; FILING DATE: 03-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/071,880
; FILING DATE: June 1, 1993
; APPLICATION NUMBER: 07/891,301
; FILING DATE: May 29, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hana Dolezalova
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.19-2 (HHD)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-1677
; TELEFAX: (415) 324-1678
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Cryptosporidium parvum
; FEATURE:
; NAME/KEY: Positions coded by nonsense codons are
; NAME/KEY: identified as Xaa.
US-08-415-751-6

Query Match      100.0%; Score 39; DB 1; Length 362;
Best Local Similarity 50.0%; Pred. No. 6.2e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 GYWXWXXW 8
DB      216 GYWLWTWN 223

RESULT 12
US-09-248-796A-15188
; Sequence 15188, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15188
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15188

Query Match      100.0%; Score 39; DB 4; Length 367;
Best Local Similarity 50.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 GYWXWXXW 8
DB      276 GYWLVDWE 283

RESULT 13
US-09-107-532A-6868
; Sequence 6868, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
```


APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6868:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...478
SEQUENCE DESCRIPTION: SEQ ID NO: 6868:
US-09-107-532A-6868

Query Match 100.0%; Score 39; DB 4; Length 478;
Best Local Similarity 50.0%; Pred. No. 8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 4;

QY 1 GYVXXXWX 8
DB 98 GYVLSAWL 105

RESULT 14
US-09-107-532A-6945
Sequence 6945, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6945:
SEQUENCE CHARACTERISTICS:

LENGTH: 492 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...492
SEQUENCE DESCRIPTION: SEQ ID NO: 6945:
US-09-107-532A-6945

Query Match 100.0%; Score 39; DB 4; Length 492;
Best Local Similarity 50.0%; Pred. No. 8.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 4;

QY 1 GYVXXXWX 8
DB 112 GYWLTCWL 119

RESULT 15
US-09-252-991A-23328
Sequence 23328, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 23328
LENGTH: 499
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23328

Query Match 100.0%; Score 39; DB 4; Length 499;
Best Local Similarity 50.0%; Pred. No. 8.4e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 4;

QY 1 GYVXXXWX 8
DB 111 GYVLSAWL 118

Search completed: January 3, 2005, 16:33:49
Job time : 23.04 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:32:45 ; Search time 80.64 Seconds
(without alignments)
35.687 Million cell updates/sec

Title: US-10-046-922-68

Perfect score: 39

Sequence: 1 GYXXXXXX 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	8	13	US-10-046-922-68
2	39	100.0	9	13	Sequence 68, Appl
3	39	100.0	10	13	Sequence 36, Appl
4	39	100.0	10	13	Sequence 34, Appl
5	39	100.0	10	13	Sequence 35, Appl
6	39	100.0	13	14	US-10-046-922-73
7	39	100.0	13	15	US-10-125-869A-73
8	39	100.0	25	14	US-10-462-262-297
9	39	100.0	29	9	US-10-280-066-476
10	39	100.0	35	9	US-09-864-761-35127
11	39	100.0	47	17	US-09-864-761-40409
12	39	100.0	57	16	US-10-425-115-287762
13	39	100.0	58	17	US-10-437-963-170197
					Sequence 308836

14	39	100.0	61	17	US-10-425-115-324628	Sequence 324628,
15	39	100.0	62	17	US-10-425-115-362818	Sequence 362818,
16	39	100.0	72	16	US-10-437-963-176036	Sequence 176036,
17	39	100.0	104	17	US-10-425-115-273234	Sequence 273234,
18	39	100.0	110	15	US-10-424-599-252989	Sequence 252989,
19	39	100.0	113	14	US-10-369-493-19264	Sequence 19264, A
20	39	100.0	116	10	US-09-910-483-1	Sequence 1, Appl
21	39	100.0	116	10	US-09-910-483-5	Sequence 5, Appl
22	39	100.0	116	10	US-09-910-483-9	Sequence 9, Appl
23	39	100.0	116	10	US-09-910-483-13	Sequence 13, Appl
24	39	100.0	116	10	US-09-910-483-17	Sequence 17, Appl
25	39	100.0	116	10	US-09-910-483-21	Sequence 21, Appl
26	39	100.0	116	10	US-09-910-483-25	Sequence 25, Appl
27	39	100.0	116	10	US-09-910-483-29	Sequence 29, Appl
28	39	100.0	116	10	US-09-910-483-33	Sequence 33, Appl
29	39	100.0	116	10	US-09-910-483-37	Sequence 37, Appl
30	39	100.0	116	10	US-09-910-483-41	Sequence 41, Appl
31	39	100.0	116	10	US-09-910-483-43	Sequence 43, Appl
32	39	100.0	122	14	US-10-447-331-6	Sequence 6, Appl
33	39	100.0	123	16	US-10-437-963-173556	Sequence 173556,
34	39	100.0	126	17	US-10-425-115-260243	Sequence 260243,
35	39	100.0	129	17	US-10-425-115-357518	Sequence 357518,
36	39	100.0	138	14	US-10-160-232-86	Sequence 86, Appl
37	39	100.0	138	14	US-10-160-232-90	Sequence 90, Appl
38	39	100.0	153	15	US-10-282-122A-68730	Sequence 68730, A
39	39	100.0	161	16	US-10-767-701-54778	Sequence 54778, A
40	39	100.0	166	15	US-10-435-114-64112	Sequence 64112, A
41	39	100.0	185	17	US-10-808-807-12	Sequence 12, Appl
42	39	100.0	199	15	US-10-424-599-267810	Sequence 267810,
43	39	100.0	227	15	US-10-282-122A-64263	Sequence 64263, A
44	39	100.0	229	17	US-10-425-115-344695	Sequence 344695,
45	39	100.0	247	15	US-10-424-599-184877	Sequence 184877,

ALIGNMENTS

RESULT 1

US-10-046-922-68
; Sequence 68, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 68
; LENGTH: 8
; TYPE: PRT
; ORGANISM: peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)..(6)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (8)..(8)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-68

Query Match 100.0%; Score 39; DB 13; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8

DB 1 GYXXXXXX 8

```
RESULT 2
US-10-046-922-36
; Sequence 36, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 9
; TYPE: PRT
; ORGANISM: peptide
US-10-046-922-36

Query Match          100.0%; Score 39; DB 13; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.5e+06;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8
DB 2 GYWDWTF 9

RESULT 3
US-10-046-922-34
; Sequence 34, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 10
; TYPE: PRT
; ORGANISM: isolated peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(1)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (10)..(10)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-34

Query Match          100.0%; Score 39; DB 13; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8
DB 2 GYWLTIWG 9

RESULT 4
US-10-046-922-35
; Sequence 35, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35
; LENGTH: 10
; TYPE: PRT
; ORGANISM: isolated peptide
US-10-046-922-35

Query Match          100.0%; Score 39; DB 13; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8
DB 2 GYWLTIWG 9

RESULT 5
US-10-046-922-73
; Sequence 73, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73
; LENGTH: 10
; TYPE: PRT
; ORGANISM: peptide library
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)..(7)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (9)..(9)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-73

Query Match          100.0%; Score 39; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8
DB 2 GYXXXXXX 9

RESULT 6
US-10-125-869A-73
; Sequence 73, Application US/10125869A
; Publication No. US20030199671A1
; GENERAL INFORMATION:
; APPLICANT: Rondon, Isaac Jesus
; APPLICANT: Wu, Qi-Long
; APPLICANT: Ley, Arthur C.
; APPLICANT: Stochl, Mark
; APPLICANT: Ransohoff, Thomas C.
; APPLICANT: Potter, M. Daniel (deceased)
; TITLE OF INVENTION: BINDING MOLECULES FOR Pc-REGION
; FILE REFERENCE: 3421.1006-001
; CURRENT APPLICATION NUMBER: US/10/125,869A
; CURRENT FILING DATE: 2002-11-19
```

; PRIOR APPLICATION NUMBER: 60/284,534
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fc region binding polypeptide
US-10-125-869A-73

Query Match 100.0%; Score 39; DB 14; Length 13;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWCNVWG 8
|||:::
Db 1 GYWCNVWG 8

RESULT 7
US-10-462-262-297
; Sequence 297, Application US/10462262
; Publication No. US20040009534A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Aaron K.
; APPLICANT: Dawson, Bruce M.
; TITLE OF INVENTION: PROTEIN ANALYSIS
; FILE REFERENCE: 10280-052001
; CURRENT APPLICATION NUMBER: US/10/462,262
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/388,642
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 297
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: immunoglobulin binding polypeptide
US-10-462-262-297

Query Match 100.0%; Score 39; DB 15; Length 13;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWCNVWG 8
|||:::
Db 1 GYWCNVWG 8

RESULT 8
US-10-280-066-476
; Sequence 476, Application US/10280066
; Publication No. US20030180718A1
; GENERAL INFORMATION:
; APPLICANT: Pillutla, Renuka C.
; APPLICANT: Brissette, Renee
; APPLICANT: Spruyt, Michael
; APPLICANT: Dedova, Olga
; APPLICANT: Blume, Arthur J.
; APPLICANT: Prendergast, John
; APPLICANT: Goldstein, Neil I.
; TITLE OF INVENTION: TARGET SPECIFIC SCREENING AND ITS USE FOR IDENTIFYING TARGET BIND
; FILE REFERENCE: 2598-4009US1
; CURRENT APPLICATION NUMBER: US/10/280,066
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 60/345,471
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 537
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 476
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Eschericia coli
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: Tie1-20C-3-D116
US-10-280-066-476

Query Match 100.0%; Score 39; DB 14; Length 25;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWCNVWG 8
|||:::
Db 6 GYWCNVWG 13

RESULT 9
US-09-864-761-35127
; Sequence 35127, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35127
; LENGTH: 29
; TYPE: PRT

```
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009503.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EST_HUMAN HIT: AA343827.1, EVALUE 1.80e+00
US-09-864-761-35127
```

```
Query Match      100.0%; Score 39; DB 9; Length 29;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GYXXXXXX 8
        |||:::|
Db      17 GYQLHWM 24
```

```
RESULT 10
US-09-864-761-40409
; Sequence 40409, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40409
; LENGTH: 35
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009503.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EST_HUMAN HIT: A1248042.1, EVALUE 9.00e-03
US-09-864-761-40409
```

```
Query Match      100.0%; Score 39; DB 9; Length 35;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GYXXXXXX 8
        |||:::|
Db      25 GYQLHWM 32
```

```
RESULT 11
US-10-425-115-287762
; Sequence 287762, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 287762
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_25533C.1.pcp
US-10-425-115-287762
```

```
Query Match      100.0%; Score 39; DB 17; Length 47;
Best Local Similarity 50.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GYXXXXXX 8
        |||:::|
Db      40 GYTTFWV 47
```

```
RESULT 12
US-10-437-963-170197
; Sequence 170197, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
```

; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 170197
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68546C.1.pep
US-10-437-963-170197

Query Match 100.0%; Score 39; DB 16; Length 57;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXXW 8
DB 19 GYWLFWMV 26

RESULT 13

US-10-425-115-308836
; Sequence 308836, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 308836
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_44726C.1.pep
US-10-425-115-308836

Query Match 100.0%; Score 39; DB 17; Length 58;
Best Local Similarity 50.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXXW 8
DB 49 GYWGASWN 56

RESULT 14

US-10-425-115-324628
; Sequence 324628, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 324628
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_59132C.1.pep
US-10-425-115-324628

Query Match 100.0%; Score 39; DB 17; Length 61;
Best Local Similarity 50.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXXW 8
DB 41 GYWKQWL 48

RESULT 15

US-10-425-115-362818
; Sequence 362818, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 362818
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(62)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_94065C.1.pep
US-10-425-115-362818

Query Match 100.0%; Score 39; DB 17; Length 62;
Best Local Similarity 50.0%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXXW 8
DB 7 GYWTDFWL 14

Search completed: January 3, 2005, 16:54:15
Job time : 81.64 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:20:13 ; Search time 17.92 Seconds
(without alignments)
42.954 Million cell updates/sec

Title: US-10-046-922-68

Perfect score: 39

Sequence: 1 GYVXXXWX 8

Scoring table: BLOSUM62DX:

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	71	2 S22905	lysis protein S -
2	39	100.0	72	2 T03190	hypothetical prote
3	39	100.0	83	2 E69903	hypothetical prote
4	39	100.0	108	2 S12193	hypothetical prote
5	39	100.0	142	2 C34903	Ig heavy chain pre
6	39	100.0	187	2 G83047	hypothetical prote
7	39	100.0	218	2 S73385	hypothetical prote
8	39	100.0	218	2 D87264	hypothetical prote
9	39	100.0	227	2 S73905	hypothetical prote
10	39	100.0	250	2 A69843	CDPdiacylglycerol-
11	39	100.0	257	2 F82733	hypothetical prote
12	39	100.0	257	2 E75325	arginine-tRNA-prot
13	39	100.0	261	2 JC5806	probable mcf prot
14	39	100.0	263	2 JC5622	aquaporin 8 - mous
15	39	100.0	271	2 F83188	aquaporin 8 - rat
16	39	100.0	271	2 JC4832	phosphatidate cyti
17	39	100.0	273	2 E95268	phosphatidate cyti
18	39	100.0	279	2 AB2307	probable ABC trans
19	39	100.0	282	2 H95869	hypothetical prote
20	39	100.0	286	2 E86900	probable sugar ABC
21	39	100.0	289	2 G72215	protein F41H10.7 [
22	39	100.0	289	2 C83886	oligopeptide ABC t
23	39	100.0	344	2 C82611	hypothetical prote
24	39	100.0	345	2 T37139	hypothetical prote
25	39	100.0	360	2 AE2047	hypothetical prote
26	39	100.0	421	2 D82500	hypothetical prote
27	39	100.0	441	2 C95307	hypothetical prote
28	39	100.0	447	2 H97146	probable transport
29	39	100.0	448	2 AB0301	siderophore/Surfac conserved hypothet

30	39	100.0	466	2 T35164	probable secreted
31	39	100.0	469	2 D70048	ABC transporter (a
32	39	100.0	472	2 E83497	probable amino aci
33	39	100.0	475	2 T46745	arginine/ornithine
34	39	100.0	482	2 JH0110	arginine/ornithine
35	39	100.0	490	2 C86879	arginine/ornithine
36	39	100.0	497	2 G86878	arginine/ornithine
37	39	100.0	508	2 C95282	probable ABC trans
38	39	100.0	517	2 A13201	hypothetical prote
39	39	100.0	519	2 S77572	oligopeptide trans
40	39	100.0	519	2 E83268	probable carbohydr
41	39	100.0	534	2 T15414	hypothetical prote
42	39	100.0	535	2 B95952	probable dipeptide
43	39	100.0	536	2 G95389	probable ABC trans
44	39	100.0	541	2 AC2392	hypothetical prote
45	39	100.0	563	2 AH2975	hypothetical prote

ALIGNMENTS

RESULT 1

S22905

lysis protein S - phase 21

C:Species: phase 21

C:Date: 09-Jun-1994 #sequence_revision 10-Nov-1995 #text_change 08-Oct-1999

C:Accession: S22905

R:Bonovich, M.T.; Young, R.

J. Bacteriol. 173, 2897-2905, 1991

A:Title: Dual start motif in two lambdaoid S genes unrelated to lambda S.

A:Reference number: S22905; MUID:91210180; PMID:2019562

A:Accession: S22905

A:Molecule type: DNA

A:Residues: 1-71 <BON>

A:Cross-references: EMBL:M65239; NID:g215466; PIDN:AAA32349.1; PID:g215467

C:Genetics:

A:Gene: S

C:Keywords: alternative initiators; cell wall lysis; transmembrane protein

F:1-71/Product: lysis protein inhibitor S107 #status predicted <MAT1>

F:4-71/Product: lysis protein S105 #status predicted <MAT2>

F:7-28/Domain: transmembrane #status predicted <TM1>

F:36-58/Domain: transmembrane #status predicted <TM2>

Query Match	100.0%	Score 39;	DB 2;	Length 71;
Best Local Similarity	50.0%;	Pred. No. 51;		
Matches	4;	Conservative	4;	Mismatches 0; Indels 0; Gaps 0;

QY 1 GYVXXXWX 8

DB 21 GYVFLQWL 28

RESULT 2

T03190

hypothetical protein 72B - rice mitochondrion

C:Species: mitochondrion Oryza sativa (rice)

C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004

C:Accession: T03190

R:Itadani, H.; Wakasugi, T.; Sugita, M.; Sugiura, M.; Nakazono, M.; Hirai, A.

Plant Cell Physiol. 35, 1239-1244, 1994

A:Title: Nucleotide sequence of a 28-kbp portion of rice mitochondria

A:Reference number: Z14841; MUID:95211382; PMID:7545979

A:Accession: T03190

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-72 <TA>

A:Cross-references: UNIPROT:Q35302; EMBL:D32052; NID:g769704; PIDN:BAA06811.1; PID:g76970

A:Experimental source: cultivar Nipponbare

C:Genetics:

A:Genome: mitochondrion

C:Keywords: mitochondrion

Query Match	100.0%	Score 39;	DB 2;	Length 72;
-------------	--------	-----------	-------	------------

Best Local Similarity 50.0%; Pred. No. 51; Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;	
QY 1 GYWXXXWX 8 : :	
Db 34 GYWSSHWI 41	
RESULT 3 E69903 hypothetical protein yodI - Bacillus subtilis C:Species: Bacillus subtilis C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004 C:Accession: E69903 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteaux C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F. Koetter, P.; Koningsgatein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A. A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis. A:Reference number: A69580; MUID:98044033; PMID:9384377 A:Accession: E69903 A>Status: preliminary; nucleic acid sequence not shown; translation not shown A:Molecule type: DNA A:Residues: 1-83 <KUN> A:Cross-references: UNIPROT:Q34654; GB:299114; GB:AL009126; NID:g2634230; PIDN:CAB13852. A:Experimental source: strain 168 C:Genetics: A:Gene: yodI	
Query Match 100.0%; Score 39; DB 2; Length 83; Best Local Similarity 50.0%; Pred. No. 58; Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;	
QY 1 GYWXXXWX 8 : :	
Db 53 GYWGYYWG 60	
RESULT 4 S12193 hypothetical protein 4 - Thiobacillus ferrooxidans plasmid pTF1 C:Species: Thiobacillus ferrooxidans C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004 C:Accession: S12193 R:Brolet, M.; Zanga, P.; Lau, P.C.K. Mol. Microbiol. 4, 1381-1391, 1990 A:Title: The mobilization and origin of transfer regions of a Thiobacillus ferrooxidans A:Reference number: S12188; MUID:91125140; PMID:2280689 A:Accession: S12193 A>Status: preliminary; translation not shown A:Molecule type: DNA A:Residues: 1-108 <DRO> A:Cross-references: UNIPROT:P20088; EMBL:X52699; NID:g48158; PIDN:CAA36930.1; PID:g48164 C:Genetics: A:Genome: plasmid pTF1	
Query Match 100.0%; Score 39; DB 2; Length 108; Best Local Similarity 50.0%; Pred. No. 74; Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;	
QY 1 GYWXXXWX 8 : :	
Db 89 GYWRSSWR 96	

RESULT 5 C34903 IG heavy chain precursor V region (5-27) - mouse C:Species: Mus musculus (house mouse) C:Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 16-Aug-1996 C:Accession: C34903 R:Bedzyk, W.D.; Herron, J.N.; Edmundson, A.B.; Voss Jr., E.W. J. Biol. Chem. 265, 133-138, 1990 A:Title: Active site structure and antigen binding properties of idiotypically cross-react A:Reference number: A34903; MUID:90094387; PMID:2104617 A:Accession: C34903 A>Status: preliminary; not compared with conceptual translation A:Molecule type: mRNA A:Residues: 1-142 <BED> C:Superfamily: immunoglobulin V region; immunoglobulin homology C:Keywords: heterotetramer; immunoglobulin F:34-119/Domain: immunoglobulin homology <IMM>	
Query Match 100.0%; Score 39; DB 2; Length 142; Best Local Similarity 50.0%; Pred. No. 95; Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;	
QY 1 GYWXXXWX 8 : :	
Db 126 GYWFAYWG 133	
RESULT 6 G83047 hypothetical protein PA4793 [imported] - Pseudomonas aeruginosa (strain PA01) C:Species: Pseudomonas aeruginosa C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004 C:Accession: G83047 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Br adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin, .; Lory, S.; Olson, M.V. Nature 406, 959-964, 2000 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathog A:Reference number: A82950; MUID:20437337; PMID:10984043 A:Accession: G83047 A>Status: preliminary A:Molecule type: DNA A:Residues: 1-187 <STO> A:Cross-references: UNIPROT:Q9HV15; GB:AE004892; GB:AE004091; NID:g9951049; PIDN:AAG0817; A:Experimental source: strain PA01 C:Genetics: A:Gene: PA4793	
Query Match 100.0%; Score 39; DB 2; Length 187; Best Local Similarity 50.0%; Pred. No. 1.2e+02; Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;	
QY 1 GYWXXXWX 8 : :	
Db 109 GYGGYYWG 116	
RESULT 7 S76385 hypothetical protein - Synechocystis sp. (strain PCC 6803) C:Species: Synechocystis sp. A:Variety: PCC 6803 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004 C:Accession: S76385 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yanada, M.; Yasuda, DNA Res. 3, 109-136, 1996 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis s. A:Reference number: S74322; MUID:97061201; PMID:8905231 A:Accession: S76385	

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <KAN>
A:Cross-references: UNIPROT:Q55705; EMBL:D54000; GB:AB001339; NID:g1001484; PIDN:BA01023
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: probable alkaline phosphatase yngC

Query Match 100.0%; Score 39; DB 2; Length 218;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8
DB 74 GYWGVRWG 81
|||||:

RESULT 8
DB7264
hypothetical protein CC0125 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: D87264
R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: D87264
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <STO>
A:Cross-references: UNIPROT:Q9ABU5; GB:AE005673; NID:g13421234; PIDN:AAK22112.1; GSPDB:C
C:Genetics:
A:Gene: CC0125

Query Match 100.0%; Score 39; DB 2; Length 218;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8
DB 139 GYWRPAWR 146
|||||:

RESULT 9
S73905
CDPdiacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase (EC 2.7.8.5) pgsA - Myc
N:Alternate names: hypothetical protein A65_orf227
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342
C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S73905
R.Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A:Reference number: S73327; MUID:97105885; PMID:8948633
A:Accession: S73905
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-227 <HIM>
A:Cross-references: UNIPROT:P75520; EMBL:AE000057; GB:U00089; NID:g1674279; PIDN:AAB9622
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetics:
A:Gene: pgsA
A:Genetic code: SGC3
C:Superfamily: CDPdiacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase
C:Keywords: transferase

Query Match 100.0%; Score 39; DB 2; Length 227;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8
DB 84 GYWARWKR 91
|||||:

RESULT 10
A69843
hypothetical protein yjba - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: A69843
R.Kunat, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero
C.; Bron, S.; Brouillet, S.; Brouillet, S.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallera
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel,
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
akeuchi, M.; Tamakoshi, A.; Tanaka, R.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: A69843
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-250 <KUN>
A:Cross-references: UNIPROT:Q31597; GB:Z99110; GB:AL009126; NID:g2633472; PIDN:CAB12998.3
A:Experimental source: strain 168
C:Genetics:
A:Gene: yjba
C:Superfamily: Bacillus subtilis hypothetical protein yjba

Query Match 100.0%; Score 39; DB 2; Length 250;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8
DB 185 GYWTEN 192
|||||:

RESULT 11
F82733
arginine-trRNA-protein transferase XF1018 [imported] - Xylella fastidiosa (strain 945c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: F82733
R.anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: F82733
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-254 <SIM>
A:Cross-references: UNIPROT:Q9PELO; GB:AE003939; GB:AE003849; NID:g9105949; PIDN:AAF83821
A:Experimental source: strain 945c
R.Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; AJ
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, D.M.; Carraro, D.M.; Carrer, H.
as-Neto, E.; Docena, C.; El-Dorri, H.; Faciniani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohne
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranee, E.E.; Laigre
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A:Authors: Martins, E.M.F.; Matekuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
submitted to GenBank, June 2000

A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1018

Query Match 100.0%; Score 39; DB 2; Length 254;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWX 8
Db 22 GYWPDRWA 29

RESULT 12
E75325
probable mcfF protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: E75325
R;White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.O.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Jenter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: E75325
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-257 <WHI>
A;Cross-references: UNIPROT:Q9RSX1; GB:AE002038; GB:AE000513; NID:g6459790; PIDN:AAF1155
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2000
A;Map position: 1

Query Match 100.0%; Score 39; DB 2; Length 257;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWX 8
Db 171 GYWLNRWA 178

RESULT 13
JC5806
aquaporin 8 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 27-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: JC5806
R;Ma, T.; Yang, B.; Verkman, A.S.
Biochem. Biophys. Res. Commun. 240, 324-328, 1997
A;Title: Cloning of a novel water and urea-permeable aquaporin from mouse expressed str
A;Reference number: JC5806; MUID:98049830; PMID:9388476
A;Accession: JC5806
A;Molecule type: mRNA
A;Residues: 1-261 <WAA>
A;Cross-references: UNIPROT:P56404; DBJ:AF018952; NID:g2353796; PIDN:AA568847.1; PID:g2
C;Comment: This protein functions as a mercurial-sensitive water channel.
C;Superfamily: lens fiber membrane major intrinsic protein
C;Keywords: glycoprotein
F;92-94/Region: NPA motif
F;210-212/Region: NPA motif
F;85,139/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 39; DB 2; Length 261;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWX 8

Db 222 GYWDFHWI 229

RESULT 14
JC5622
aquaporin 8 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: JC5622
R;Fshibashi, K.; Kuwahara, M.; Kageyama, Y.; Tohaaka, A.; Marumo, F.; Sasaki, S.
Biochem. Biophys. Res. Commun. 237, 714-718, 1997
A;Title: Cloning and functional expression of a second new aquaporin abundantly expressed
A;Reference number: JC5622; MUID:97445104; PMID:9299432
A;Contents: Testis
A;Accession: JC5622
A;Molecule type: mRNA
A;Residues: 1-263 <ISH>
A;Cross-references: UNIPROT:P56405; DBJ:AB005547; NID:g2346967; PIDN:BAA21918.1; PID:g2;
C;Comment: This protein is a water channel protein which plays a role in the regulation
C;Superfamily: lens fiber membrane major intrinsic protein
C;Keywords: glycoprotein
F;39-59/Domain: transmembrane #status predicted <TM1>
F;65-94/Domain: transmembrane #status predicted <TM2>
F;94-96/Region: NPA motif
F;109-130/Domain: transmembrane #status predicted <TM3>
F;159-179/Domain: transmembrane #status predicted <TM4>
F;182-204/Domain: transmembrane #status predicted <TM5>
F;212-214/Region: NPA motif
F;231-250/Domain: transmembrane #status predicted <TM6>
F;141/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 39; DB 2; Length 263;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWX 8
Db 224 GYWDFHWI 231

RESULT 15
F83188
phosphatidate cytidyltransferase PA3651 [imported] - Pseudomonas aeruginosa (strain PA
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: F83188
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: AB2950; MUID:20437337; PMID:10984043
A;Accession: F83188
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-271 <STO>
A;Cross-references: UNIPROT:Q59640; GB:AE004785; GB:AE004091; NID:g9949809; PIDN:AA0703;
A;Experimental source: strain PA01
C;Genetics:
A;Gene: cdsA; PA3651
C;Superfamily: phosphatidate cytidyltransferase

Query Match 100.0%; Score 39; DB 2; Length 271;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWX 8
Db 102 GYWGGRWR 109

Search completed: January 3, 2005, 16:34:52

Job time : 17.92 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 16:12:00 ; Search time 97.28 Seconds
(without alignments)
47.317 Million cell updates/sec

Title: US-10-046-922-68
Perfect score: 39
Sequence: 1 GYVXXXXW 8

Scoring table: BL0SUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:
1: uniprot_sprot:
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	30	Q7UDB7	Q7UDB7 shigella fl
2	39	100.0	49	Q9EVP1	Q9EVP1 escherichia
3	39	100.0	71	VLXS_BPP21	P27360 bacteriophage
4	39	100.0	71	Q7C2J0	Q7C2J0 shigella fl
5	39	100.0	71	Q9FCW3	Q9FCW3 escherichia
6	39	100.0	71	Q7UDP0	Q7UDP0 shigella fl
7	39	100.0	71	Q83ML2	Q83ML2 shigella fl
8	39	100.0	72	Q83S57	Q83S57 shigella fl
9	39	100.0	72	Q35302	Q35302 oryza sativ
10	39	100.0	83	YODI_BACSU	Q34654 bacillus su
11	39	100.0	89	Q9SS05	Q9SS05 drosophila
12	39	100.0	102	Q6IIL4	Q6IIL4 drosophila
13	39	100.0	108	YML2_THIFE	P20088 thiobacilli
14	39	100.0	120	Q728A6	Q728A6 desulfovibr
15	39	100.0	120	AAS97170	AAS97170 desulfovi
16	39	100.0	122	Q72D02	Q72D02 desulfovibr
17	39	100.0	122	AAS95609	AAS95609 desulfovi
18	39	100.0	124	Q7U7V6	Q7U7V6 synchococc
19	39	100.0	128	Q8MK57	Q8MK57 bos taurus
20	39	100.0	130	Q7U395	Q7U395 prochloroco
21	39	100.0	130	Q7VBS3	Q7VBS3 prochloroco
22	39	100.0	135	Q7TUW7	Q7TUW7 prochloroco
23	39	100.0	160	Q6NF17	Q6NF17 corynebacte
24	39	100.0	160	CAES0615	CAES0615 corynebac
25	39	100.0	187	Q9HV15	Q9HV15 pseudomonas
26	39	100.0	189	Q880B8	Q880B8 pseudomonas
27	39	100.0	204	Q7W0P5	Q7W0P5 bordetella
28	39	100.0	204	Q7W3F9	Q7W3F9 bordetella
29	39	100.0	204	Q7WES9	Q7WES9 bordetella
30	39	100.0	212	Q7WL18	Q7WL18 bordetella
31	39	100.0	213	Q9KY37	Q9KY37 streptomyce

32 39 100.0 214 2 Q8A8U4 Q8A8U4 bacteroides
33 39 100.0 218 1 Y232_SINY3 Y232_SINY3 synchococyst
34 39 100.0 218 2 Q7VZM7 Q7VZM7 bordetella
35 39 100.0 218 2 Q9ABU5 Q9ABU5 caulobacter
36 39 100.0 224 2 Q8S486 Q8S486 zea mays (m
37 39 100.0 227 1 FGSA_MYCPN FGSA_MYCPN mycoplasma
38 39 100.0 228 2 Q7ZKR5 Q7ZKR5 thermus the
39 39 100.0 228 2 AAS80801 AAS80801 thermus t
40 39 100.0 235 2 Q7W7N0 Q7W7N0 bordetella
41 39 100.0 236 2 Q9WGW9 Q9WGW9 human immun
42 39 100.0 243 2 Q8CX61 Q8CX61 arabidopsis
43 39 100.0 250 2 Q31597 Q31597 bacillus su
44 39 100.0 253 2 Q32816 Q32816 lactococcus
45 39 100.0 254 1 ATE_XYLF ATE_XYLF xylella fag

ALIGNMENTS

RESULT 1
Q7UDB7 PRELIMINARY; PRT; 30 AA.
AC Q7UDB7;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical bacteriophage protein.
GN Name=vbcR; OrderedLocusNames=S0714;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AB016980; AAP16193.1; -.
DR InterPro: IPR007054; Lysis_S.
DR Pfam: PF04971; Lysis_S; 1.
KW Hypothetical protein.

Query Match 100.0%; Score 39; DB 2; Length 30;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYVXXXXW 8
DB 21 GYVFLQWL 28

RESULT 2
Q9EVP1 PRELIMINARY; PRT; 49 AA.
AC Q9EVP1;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE S protein (fragment).
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H.1.8.;

```
RX MEDLINE=20407286; PubMed=10948037;
RA Unkneir A., Schmidt H.;
RT "Structural analysis of phage-borne stx genes and their flanking
RT sequences in shiga toxin-producing Escherichia coli and Shigella
RT dysenteriae type 1 strains.";
RL Infect. Immun. 68:4856-4864(2000).
DR ENBL; AJ271139; CAC05573.1; -.
DR InterPro; IPR007054; Lysis_S.
DR Pfam; PF04971; Lysis_S; 1.
FT NON TER 49
SQ SEQUENCE 49 AA; 5227 MW; 0B6914DD9AE25E00 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 49;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXX 8
DB 21 GYWFLQWL 28

RESULT 3
VLVS BPP21
ID VLVS_BPP21 STANDARD; PRT; 71 AA.
AC P27320;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lysis protein S.
GN Name=S;
OS Bacteriophage P21 (Bacteriophage 21).
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10711;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91210180; PubMed=2019562;
RA Bonovich M.T., Young R.;
RT "Dual start motif in two lambdoid S genes unrelated to lambda S.";
RL J. Bacteriol. 173:2897-2905(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; M65239; AAA32349.1; -.
DR InterPro; IPR007054; Lysis_S.
DR Pfam; PF04971; Lysis_S; 1.
KW Phage lysis protein.
SQ SEQUENCE 71 AA; 7893 MW; 8690A8F25234A3E2 CRC64;

Query Match 100.0%; Score 39; DB 1; Length 71;
Best Local Similarity 50.0%; Pred. No. 4e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXX 8
DB 21 GYWFLQWL 28

RESULT 4
Q7C2J0
ID Q7C2J0 PRELIMINARY; PRT; 71 AA.
AC Q7C2J0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative S protein.
GN OrderedLocusNames=S0731;
```

```
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=24527;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR ENBL; AE016980; AAP16205.1; -.
DR InterPro; IPR007054; Lysis_S.
DR Pfam; PF04971; Lysis_S; 1.
SQ SEQUENCE 71 AA; 7314 MW; 10CE1C485234AE99 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 71;
Best Local Similarity 50.0%; Pred. No. 4e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXX 8
DB 21 GYWFLQWL 28

RESULT 5
Q9FCW3
ID Q9FCW3 PRELIMINARY; PRT; 71 AA.
AC Q9FCW3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE S protein.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T4/97;
RX MEDLINE=20407286; PubMed=10948097;
RA Unkneir A., Schmidt H.;
RT "Structural analysis of phage-borne stx genes and their flanking
RT sequences in shiga toxin-producing Escherichia coli and Shigella
RT dysenteriae type 1 strains.";
RL Infect. Immun. 68:4856-4864(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=T4/97;
RA Unkneir A., Karch H., Schmidt H.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=T4/97;
RA Schmidt H., Scheef J., Morabito S., Caprioli A., Wieler L., Karch H.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR ENBL; AJ270998; CAC05565.1; -.
DR InterPro; IPR007054; Lysis_S.
DR Pfam; PF04971; Lysis_S; 1.
SQ SEQUENCE 71 AA; 7323 MW; 9B4D68F25220B7E2 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 71;
Best Local Similarity 50.0%; Pred. No. 4e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXX 8
DB 21 GYWFLQWL 28
```



```
RESULT 6
Q7UDP0
ID Q7UDP0 PRELIMINARY; PRT; 71 AA.
AC Q7UDP0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-NAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lysis protein S.
GN OrderedLocusNames=S0231;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; A016978; AAP1579.1; -.
DR InterPro; IPR007054; Lysis_S.
DR Pfam; PF04971; Lysis_S; 1.
DR Complete proteome.
SQ SEQUENCE 71 AA; 7861 MW; AB92BAF25234BC15 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 71;
Best Local Similarity 50.0%; Pred. No. 4e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXXW 8
Db 21 GYWFQWL 28

RESULT 7
Q83ML2
ID Q83ML2 PRELIMINARY; PRT; 71 AA.
AC Q83ML2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE LYSIS PROTEIN S.
GN OrderedLocusNames=SF2038;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL; A015099; AAN42324.1; -.
DR InterPro; IPR007054; Lysis_S.
DR Pfam; PF04971; Lysis_S; 1.
DR Complete proteome.
SQ SEQUENCE 71 AA; 7865 MW; AB90A8F25234A3P5 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 71;
Best Local Similarity 50.0%; Pred. No. 4e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXXW 8
Db 21 GYWFQWL 28

RESULT 8
Q83S57
ID Q83S57 PRELIMINARY; PRT; 71 AA.
AC Q83S57;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Lysis protein S.
GN OrderedLocusNames=SF0689;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL; A015099; AAN42324.1; -.
DR InterPro; IPR007054; Lysis_S.
DR Pfam; PF04971; Lysis_S; 1.
DR Complete proteome.
SQ SEQUENCE 71 AA; 7914 MW; 10CE1C485234AB99 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 71;
Best Local Similarity 50.0%; Pred. No. 4e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXXW 8
Db 21 GYWFQWL 28

RESULT 9
Q35302
ID Q35302 PRELIMINARY; PRT; 72 AA.
AC Q35302;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ORF72B.
OS Oryza sativa (japonica cultivar-group).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC Baev A.A., Dzhmagalliev E.B., Lyubomirskaya N.V., Mizrokhi L.Y.,
RA Il'in Y.V.;
RT "Structure of long and short copies of the mobile dispersed gene MDG3
RT of Drosophila melanogaster."
RL Dokl. Akad. Nauk SSSR 282:1483-1486(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95308541; PubMed=7788722;
RA Nakazono M., Itadani H., Wakasugi T., Tsutsumi N., Sugiura M.,
RA Hirai A.;
RT "The rps3-rpl16-nad3-rps12 gene cluster in rice mitochondrial DNA is
```

```

RT transcribed from alternative promoters.";
RL Curr. Genet. 27:184-189(1995).
RN [3]
RX MEDLINE=95211382; PubMed=7545979;
RA Itadani H., Wakasugi T., Sugita M., Sugiura M., Nakazono M., Hirai A.;
RT "Nucleotide sequence of a 28-kbp portion of rice mitochondrial DNA:
RT the existence of many sequences that correspond to parts of
RT mitochondrial genes in intergenic regions.";
RL Plant Cell Physiol. 35:1239-1244(1994).
DR EMBL; D32052; BAA06811.1; -.
DR PIR; T03190; T03190.
DR Gramene; Q35302; -.
DR GO; GO:0005739; C-mitochondrion; IEA.
KW Mitochondrion.
SQ
-----
Query Match 100.0%; Score 39; DB 2; Length 72;
Best Local Similarity 50.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXXW 8
Db |||:::|
34 GYWSHWI 41

RESULT 10
YODI_BACSU
ID YODI_BACSU STANDARD; PRT; 83 AA.
AC Q34654;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical protein yodi.
GN Name=yodi; Synonyms=yolA; OrderedLocusNames=BSU19610;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RA Wambutt R., Wedler H., Lapidus A., Sorokin A., Ehrlich S.D.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
RT the odhA and aspC loci cloned in a yeast artificial chromosome.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Ghim S.-Y., Jeong Y.-M., Choi S.-K., Park S.-H.;
RT "Sequence analysis of the 30 kb region (182') of the Bacillus subtilis
RT chromosome containing the cge cluster.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunes F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Broutillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.-K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N.,
RA Ghim S.-Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA Jones L.-M., Joris B., Karanata D., Kasahara Y., Klaerr-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningstein G., Krogh S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,

```

```

RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadate Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weizenegger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL Nature 390:249-256(1997).
CC
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
DR EMBL; AF015775; AAB72056.1; -.
DR EMBL; AF006665; AAB81166.1; -.
DR EMBL; Z99114; CAB13852.1; -.
DR PIR; E69903; E69903.
DR Subtilist; BG13537; yodi.
DR InterPro; IPR008991; Transl SH3 like.
KW Complete proteome; Hypothetical protein; Transmembrane.
FT TRANSMEM 58 80 Potential.
SQ SEQUENCE 83 AA; 9194 MW; 99F58EA2F0F36A43 CRC64;

Query Match 100.0%; Score 39; DB 1; Length 83;
Best Local Similarity 50.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXXW 8
Db |||:::|
53 GYWGYYWG 60

RESULT 11
Q95S05 PRELIMINARY; PRT; 89 AA.
AC Q95S05;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE HL03793p.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY061016; AAL28564.1; -.
DR FlyBase; FBgn0047239; BcDNA:HL03793.
SQ SEQUENCE 89 AA; 10063 MW; 36CB86917DF80B9D CRC64;

Query Match 100.0%; Score 39; DB 2; Length 89;
Best Local Similarity 50.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXXW 8

```

```
DB 70 GYWCNWE 77
||||:|
SQ SEQUENCE 108 AA; 12335 MW; A8E67717C109A57E CRC64;
Query Match 100.0%; Score 39; DB 1; Length 108;
Best Local Similarity 50.0%; Pred. No. 5.8e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

RESULT 12
Q611L4 PRELIMINARY; PRT; 102 AA.
ID Q611L4
AC Q611L4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HDC17696.
GN ORFNAMES=HDC17696;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14709175;
RA Hild M., Beckmann B., Haas S., Koch B., Solovjev V., Busold C.,
RA Feilenberg K., Boutsos M., Vingron M., Sauer F., Honeisei J., Paro R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome.";
RL Genome Biol. 5:R3-R3(2003).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK003052; DAA03252.1; -.
SQ SEQUENCE 102 AA; 11777 MW; CD4044EC9325CP46 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 102;
Best Local Similarity 50.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXX 8
||||:|
DB 73 GYWIQWS 80

RESULT 13
YML2_THIFE STANDARD; PRT; 108 AA.
ID YML2_THIFE
AC P20088;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 12.3 kDa protein in mobil 3' region (ORF 4).
OS Thiobacillus ferrooxidans.
OG Plasmid pTF1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Acidithiobacillales;
OC Acidithiobacillaceae; Acidithiobacillus.
OX NCBI_TaxID=920;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 33020;
RX MEDLINE=91125140; PubMed=2280689;
RA Drollet M., Zanga P., Lau P.C.K.;
RT "The mobilization and origin of transfer regions of a Thiobacillus
RT ferrooxidans plasmid: relatedness to plasmids RSP1010 and pSCI101.";
RL Mol. Microbiol. 4:1381-1391(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52699; CAA36930.1; -.
DR PIR; S12193; S12193.
KW Hypothetical protein; Plasmid.

SQ SEQUENCE 108 AA; 12335 MW; A8E67717C109A57E CRC64;
Query Match 100.0%; Score 39; DB 1; Length 108;
Best Local Similarity 50.0%; Pred. No. 5.8e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXX 8
||||:|
DB 89 GYWRGSR 96

RESULT 14
Q728A6 PRELIMINARY; PRT; 120 AA.
ID Q728A6
AC Q728A6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Lipoprotein, putative.
GN OrderedLocusNames=DVU2698;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 9303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Daviden T.M., Zafar N., Zhou L., Radune D.,
RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldlyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AS017318; AAS97170.1; -.
DR TIGR; DVU2698; -.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 120 AA; 13418 MW; CD3581657D76E183 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 120;
Best Local Similarity 50.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXWXX 8
||||:|
DB 54 GYWDRA 61

RESULT 15
AAS97170 PRELIMINARY; PRT; 120 AA.
ID AAS97170
AC AAS97170;
DT 26-APR-2004 (TrEMBLrel. 27, Created)
DT 26-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 11-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Lipoprotein, putative.
GN DVU2698.
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 9303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15077118;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
```

RA Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,
 RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
 RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
 RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
 RT *Desulfovibrio vulgaris* Hildenborough.",
 RL Nat. Biotechnol. 22:554-559(2004).
 DR EMBL; AS017318; AAS97170.1; -.
 DR TIGR; DVU2698; -.
 KW Lipoprotein.
 SQ SEQUENCE 120 AA; 13418 MW; CD3581657D76E183 CRC64;

 Query Match 100.0%; Score 39; DB 2; Length 120;
 Best Local Similarity 50.0%; Pred.No. 6.3e+02;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

 QY 1 GYWXWX 8
 Db 54 GYWIDRWA 61
 |||::|:
 |||::|:

Search completed: January 3, 2005, 16:32:33
 Job time : 97.28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 17:22:59 ; Search time 154 Seconds
(without alignments)
18.635 Million cell updates/sec

Title: US-10-046-922-68

Perfect score: 39

Sequence: 1 GYXXXXWX 8

Scoring table: BLOSUM62DX-4
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 153769

Minimum DB seq length: 0

Maximum DB seq length: 48

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	8	5	Abp53965 VEGFR-3 b
2	38	97.4	7	3	Aay76794 Somatosta
3	38	97.4	7	5	Abp53964
4	38	97.4	7	5	Abp53418 Backbone
5	33	84.6	7	8	Adj25834 Tyrosine
6	33	84.6	8	2	Aaw27066 Galactana
7	33	84.6	8	4	Abp24270 HIV A24 m
8	33	84.6	8	4	Abp15891 HIV A24 s
9	33	84.6	8	4	Abp15892
10	33	84.6	8	4	Abp24235
11	32	82.1	6	2	Aar93713 Cyclo(-ty
12	32	82.1	6	5	Aau83934 Tyrosine
13	32	82.1	7	2	Aar08140 Neurokin
14	32	82.1	7	3	Aay76792 Somatosta
15	32	82.1	7	5	Abp53416 Backbone
16	32	82.1	8	1	Aap61465 Cyclic oc
17	32	82.1	8	1	Aap90981 Water-ins
18	32	82.1	8	2	Aar14237 Somatosta
19	32	82.1	8	2	Aar15362 Somatosta
20	32	82.1	8	2	Aar10505 Pamote o
21	32	82.1	8	2	Aar26369 Somatosta
22	32	82.1	8	2	Aar27185 Somatosta
23	32	82.1	8	2	Aar27187 Somatosta
24	32	82.1	8	2	Aar27197 Somatosta
25	32	82.1	8	2	Aar23952 LH-RH, so

26	32	82.1	8	2	AAR40822	Aar40822 Octapepti
27	32	82.1	8	2	AAR41502	Aar41502 Somatosta
28	32	82.1	8	2	AAR32766	Aar32766 Lanthioni
29	32	82.1	8	2	AAR31513	Aar31513 Somatosta
30	32	82.1	8	2	AAR31511	Aar31511 Somatosta
31	32	82.1	8	2	AAR42653	Aar42653 Somatosta
32	32	82.1	8	2	AAR56781	Aar56781 Somatosta
33	32	82.1	8	2	AAR76207	Aar76207 Somatosta
34	32	82.1	8	2	AAR76205	Aar76205 Somatosta
35	32	82.1	8	2	AAR85570	Aar85570 Somatosta
36	32	82.1	8	2	AAI02418	Aay02418 Somatosta
37	32	82.1	8	2	Aaw18454	Aaw18454 Somatosta
38	32	82.1	8	2	AAW51865	Aay1865 Somatosta
39	32	82.1	8	2	AAI22044	Aay22044 Somatosta
40	32	82.1	8	2	AAI18229	Aay18229 Somatosta
41	32	82.1	8	2	AAW45740	Aaw45740 Somatosta
42	32	82.1	8	2	AAW97185	Aaw97185 Somatosta
43	32	82.1	8	2	ADH35407	Adh35407 Human som
44	32	82.1	8	2	ADH68008	Adh68008 Somatosta
45	32	82.1	8	2	ADK11160	Adk11160 Somatosta

ALIGNMENTS

RESULT 1

ABP53965
ID ABP53965 standard; peptide; 8 AA.

XX ABP53965;

XX AC

DT 09-JAN-2003 (first entry)

DE VEGFR-3 binding peptide SEQ ID NO:68.

XX

Vascular endothelial growth factor receptor 3 inhibitor; VEGFR-3; angiogenesis; lymphangiogenesis; vascular endothelial growth factor; cytotactic; hepatotropic; antiinflammatory; hypotensive; antidiabetic; KW vulnary; cell surface receptor; cancer; neovascularisation; KW liver disease; hypertension; post-trauma; chronic hepatitis; haemangioma; KW diabetes; PDGF; platelet derived growth factor.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 4..6 /note= "X is any amino acid"

FT Misc-difference 8 /note= "any amino acid"

WO200257299-A2.

25-JUL-2002.

16-JAN-2002; 2002WO-IB000099.

17-JAN-2001; 2001US-0262476P.

(LUDW-) LUDWIG INST CANCER RES.

(LICN) LICENTIA LTD.

PI Alitalo K, Koivunen E, Kubo H;

XX WPI; 2002-691521/74.

XX New isolated peptide that inhibits VEGF-C and VEGF-D, useful for diagnosing, evaluating, treating disorders mediated by VEGFR-3 activity, such as cancer and diseases of neovascularization.

PS Claim 22; Page 81; 149pp; English.

CC The present invention describes an isolated peptide (I) that binds to and

XX PS Claim 21; Page 81; 149pp; English.

XX CC The present invention describes an isolated peptide (I) that binds to and

CC inhibits vascular endothelial growth factor receptor 3 (VEGFR-3). (I)

CC have cytostatic, hepatotropic, antiinflammatory, hypotensive,

CC antidiabetic and vulnerary activities, and can be used in gene therapy.

CC Compositions and methods from the present invention are useful for

CC diagnosing, evaluating and treating disorders mediated by the activity of

CC the cell surface receptor VEGFR-3 such as cancer, e.g. brain, lung,

CC liver, spleen, kidney, lymph node, small intestine, blood cells,

CC pancreas, colon, stomach, breast, endometrium, prostate, testicle, ovary,

CC skin, head and neck, oesophagus, bone, marrow or blood, and diseases of

CC neovascularisation, e.g. liver diseases, hypertension, post-trauma,

CC chronic hepatitis, haemangiomas and diabetes. The present sequence

CC represents a specifically claimed VEGFR-3 binding peptide from the

XX present invention

XX SQ Sequence 7 AA;

Query Match 97.4%; Score 38; DB 5; Length 7;

Best Local Similarity 100.0%; Pred. No. 1.7e+06; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7

DB 1 GYWXXXW 7

|||||

RESULT 4

ABP53418

ID ABP53418 standard; peptide; 7 AA.

XX AC ABP53418;

XX DT 19-NOV-2002 (first entry)

XX DE Backbone cyclised somatostatin analogue PTR 3181.

XX KW Backbone cyclised somatostatin analogue; somatostatin; SRIP; analgesic;

KW somatostatin release inhibiting factor; somatostatin receptor subtype;

KW synthesis; antiarteriosclerotic; immunosuppressive; cytostatic; cancer;

KW antidiabetic; antiinflammatory; somatostatin receptor ligand;

KW atherosclerosis; autoimmune disease; diabetic-associated complication;

KW endocrine disorder; inflammation; gastrointestinal disorder; restenosis;

XX pancreatitis; post-surgical pain.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminally modified with Fmoc

FT (fluorenylmethoxycarbonyl)"

FT Misc-difference 3

FT Modified-site 7 /note= "D form residue"

FT /note= "amidated"

PN US2002052315-A1.

XX PD 02-MAY-2002.

XX PF 13-DEC-2000; 2000US-00734583.

XX PR 19-JUN-1998; 98US-00100360.

PR 02-DEC-1998; 98US-00203389.

PR 15-JUN-1999; 99WO-IL000329.

XX (HORN/) HORNIK V.

PA (AFAR/) AFARGAN M M.

PA (GELL/) GELLERMAN G.

XX Hornik V, Afargan MM, Gellerman G;

XX WPI; 2002-681319/73.

XX PT New backbone cyclized somatostatin analogs are e.g. useful in the

PT treatment of atherosclerosis, autoimmune diseases and cancers.

XX Example 12; Page 21; 30pp; English.

XX CC The present invention describes backbone cyclised somatostatin analogues

CC (I) that incorporate at least one building unit containing one nitrogen

CC atom of the peptide backbone connected to a bridging group (comprising an

CC amide, thioether, thioester or disulfide) where at least one building

CC unit is connected via the bridging group to form a cyclic structure with

CC a moiety selected from the group consisting of a second building unit,

CC the side chain of an amino acid residue of the sequence or the N-terminal

CC amino acid residue. (I) has antiarteriosclerotic, immunosuppressive,

CC cytostatic, antidiabetic, antiinflammatory and analgesic activities, and

CC can be used as a somatostatin receptor ligand. (I) are useful in the

CC treatment of atherosclerosis, autoimmune diseases, cancers, diabetic-

CC associated complications, endocrine disorders, inflammation,

CC gastrointestinal disorders, pancreatitis, post-surgical pain, and

CC restenosis. (I) can also be used in the diagnosis of cancer, by imaging

CC the existence of metastases, it being labeled with a detectable probe.

CC The present sequence represents a backbone cyclised somatostatin analogue

CC from the present invention

XX SQ Sequence 7 AA;

Query Match 97.4%; Score 38; DB 5; Length 7;

Best Local Similarity 57.1%; Pred. No. 1.7e+06;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7

DB 1 GYWKVCW 7

|||||

RESULT 5

ADJ25834

ID ADJ25834 standard; peptide; 7 AA.

XX AC ADJ25834;

XX DT 20-MAY-2004 (first entry)

XX DE Tyrosine tRNA synthetase binding peptide group 2 motif.

XX KW ligand identification; peptide library;

KW complementary combinatorial library; tyrosine tRNA synthetase.

XX OS Synthetic.

XX PN US6617114-B1.

XX PD 09-SEP-2003.

XX PF 30-APR-1998; 98US-00069827.

XX PR 31-OCT-1996; 96US-00740671.

PR 31-OCT-1997; 97WO-US019638.

PR 31-MAR-1998; 98US-00050359.

XX (KARO-) KARO BIO AB.

XX Fowlkes DM, Kay BK, Frelinger JA, Hyde-Deruysscher RP;

WPI; 2004-068186/07.

XX Identification of ligand that can mediate biological activity of target

PT protein, comprises screening first combinatorial library having first

PT member ligands for binding to target protein to identify target-binding

PT ligand(s).

PS Example 5; SEQ ID NO 94; 98pp; English.

XX The invention relates to a method of identifying a ligand that can mediate the biological activity of target protein via inhibition of the binding of target protein to a binding partner ligand comprising screening first combinatorial library having first member ligands for binding to target protein to identify target-binding ligand(s). The method is useful for identifying ligands that can mediate the biological activity of target proteins via inhibition of the binding of target protein to a binding partner ligand. The invention does not require that the natural binding partner be used as reagent. The need for the natural binding partner is obviated with the use of complementary combinatorial libraries. The present sequence is used in the exemplification of the present invention.

XX

XX Sequence 7 AA;

Query Match 84.6%; Score 33; DB 8; Length 7;
 Best Local Similarity 42.9%; Pred. No. 1.7e+06;
 Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXWX 8
 ||::||
 1 YWPDWG 7

Db

RESULT 6
 AAW27066
 ID AAW27066 standard; peptide; 8 AA.
 AC AAW27066;
 XX 17-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 28-FEB-1998 (first entry)
 XX Galactanase peptide motif.
 DE Galactanase; fruit juice; vegetable juice; wine; pectin; depectinisation;
 XX animal feed; foodstuff.
 KW Corynascus heterochallicus; strain CBS 117.65.
 XX Humicola insolens; strain DSM 1800.
 OS

XX WO9732014-A1.
 XX 04-SEP-1997.
 XX 28-FEB-1997; 97WO-DK000092.
 XX 01-MAR-1996; 96DK-00000233.
 PR 01-MAR-1996; 96DK-00000235.
 XX (NOVO) NOVO-NORDISK AS.
 PA

XX Kofod LV, Kauppinen MS, Andersen LN, Clausen IG, Muellertz A;
 PI WPI; 1997-448686/41.
 DR Fungal galactanase and related DNA - useful in animal feed industry and fruit juice depectinisation.
 XX

XX Claim 7; Page 58; 67pp; English.

XX This peptide comprises a motif located at amino acid residues 312-319 of the galactanase of Myceliophthora thermophila CBS 117.65 (see AAW27063) and at amino acid residues 311-318 of the galactanase of Humicola insolens DSM 1800 (see AAW27064). Another motif (see AAW27065) has also been identified. These motifs are characteristic of galactanases of fungi of the order Sordariales. PCR primers (see AAR85060-61) based on the motifs can be used to identify galactanase enzymes in Sordariales fungi. Such enzymes have an optimum pH above 5.8 and are useful in the food and feed industries, as well as in wine and juice processing. (Updated on 25-

CC MAR-2003 to correct PI field.) (Updated on 17-OCT-2003 to standardise OS field)

CC

XX Sequence 8 AA;

Query Match 84.6%; Score 33; DB 2; Length 8;
 Best Local Similarity 42.9%; Pred. No. 1.7e+06;
 Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXWX 8
 ||::||
 2 YWEPAMI 8

Db

RESULT 7
 ABP24270
 ID ABP24270 standard; peptide; 8 AA.
 XX ABP24270;
 AC ABP24270;
 XX 11-SEP-2003 (revised)
 DT 15-JUL-2002 (first entry)
 XX HIV A24 motif pol peptide #78.
 DE

XX HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
 KW vaccine; HIV infection; immunisation; virucide.
 XX

OS Human immunodeficiency virus 1.
 XX WO200124810-A1.
 XX 12-APR-2001.
 PD 05-OCT-2000; 2000WO-US027766.
 XX 05-OCT-1999; 99US-00412863.
 PR (EPIM-) EPIMMUNE INC.
 XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
 PI Baker DM, Celis E, Kubo RT, Grey HM;
 XX WPI; 2001-354887/37.
 DR Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1) peptide groups, useful for vaccinating against HIV-1.
 XX Claim 32; Page 366; 448pp; English.

XX The present invention describes a composition (I) comprising a prepared human immunodeficiency virus-1 (HIV-1) group comprising an amino acid sequence selected from 51 defined amino acid sequences (ABL25347 to ABP25397). (I) has virucide activity and can be used in vaccines. (I) may be used for immunising subjects against HIV-1 infections. The use of group-based vaccines has several advantages over traditional vaccines, particularly when compared to the use of whole antigens in vaccine compositions. There is evidence that the immune response to whole antigens is directed largely toward variable regions of the antigen, allowing for immune escape due to mutations. The groups for inclusion in an group-based vaccine may be selected from conserved regions of viral or tumour-associated antigens, which therefore reduces the likelihood of escape mutants. Furthermore, immunosuppressive groups that may be present in whole antigens can be avoided with the use of group-based vaccines. An additional advantage of an group-based vaccine approach is the ability to combine selected groups (CTL and HTL), and further, to modify the composition of the groups, achieving, for example, enhanced immunogenicity. Accordingly, the immune response can be modulated, as appropriate, for the target disease. Similar engineering of the response is not possible with traditional approaches. ABP1501 to ABP25412 represent peptide sequences used in the exemplification of the present invention. (Updated on 11-SEP-2003 to standardise OS field)

Best Local Similarity 42.9%; Pred. No. 1.7e+06;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXWX 8
||::|:
Db 2 YWQATWI 8

RESULT 10
ABP24235
ID ABP24235 standard; peptide; 8 AA.

XX AC ABP24235;
XX DT 11-SEP-2003 (revised)
XX DT 15-JUL-2002 (first entry)
XX DE HIV A24 motif pol peptide #43.
XX KW HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
XX KW vif; tat; cytototoxic T lymphocyte; CTL; immune response; epitope; antigen;
XX KW vaccine; HIV infection; immunisation; virucide.

XX OS Human immunodeficiency virus 1.
XX PN WO200124810-A1.

XX PD 12-APR-2001.

XX PF 05-OCT-2000; 2000WO-US027766.

XX PR 05-OCT-1999; 99US-00412863.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Celis E, Kubo RT, Grey HM;

XX DR WPI; 2001-354887/37.

XX PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
XX PT peptide groups, useful for vaccinating against HIV-1.

XX PS Claim 32; Page 366; 448pp; English.

XX CC The present invention describes a composition (I) comprising a prepared
XX CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
XX CC sequence selected from 51 defined amino acid sequences (ABP25347 to
XX CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
XX CC be used for immunising subjects against HIV-1 infections. The use of
XX CC group-based vaccines has several advantages over traditional vaccines,
XX CC particularly when compared to the use of whole antigens in vaccine
XX CC compositions. There is evidence that the immune response to whole
XX CC antigens is directed largely toward variable regions of the antigen,
XX CC allowing for immune escape due to mutations. The groups for inclusion in
XX CC an group-based vaccine may be selected from conserved regions of viral or
XX CC tumour-associated antigens, which therefore reduces the likelihood of
XX CC escape mutants. Furthermore, immunosuppressive groups that may be present
XX CC in whole antigens can be avoided with the use of group-based vaccines. An
XX CC additional advantage of an group-based vaccine approach is the ability to
XX CC combine selected groups (CTL and HTL), and further, to modify the
XX CC composition of the groups, achieving, for example, enhanced
XX CC immunogenicity. Accordingly, the immune response can be modulated, as
XX CC appropriate, for the target disease. Similar engineering of the response
XX CC is not possible with traditional approaches. ABP11501 to ABP25412
XX CC represent peptide sequences used in the exemplification of the present
XX CC invention. (Updated on 11-SEP-2003 to standardise OS field)

XX SQ Sequence 8 AA;

Query Match 84.6%; Score 33; DB 4; Length 8;
Best Local Similarity 42.9%; Pred. No. 1.7e+06;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXWX 8
||::|:
Db 2 YWQATWI 8

RESULT 11
AAR93713
ID AAR93713 standard; peptide; 6 AA.

XX AC AAR93713;

XX DT 10-MAY-1996 (first entry)

XX DE Cyclo[-Tyr-trp-Leu-Arg-Gly-Trp-].

XX KW neurokinin A antagonist; tachykinin; respiratory disease; asthma;
XX KW analgesic; cyclic.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Modified-site 1

XX FT /note= "not an N-terminal amino acid, but condensed with
Trp(6) to form a cyclic peptide"

XX FT Misc-difference 2

XX FT /note= "D-form residue"

XX FT Modified-site 6

XX FT /note= "not a C-terminal amino acid, but condensed with
Tyr(1) to form a cyclic peptide"

XX PN WO9521187-A1.

XX PD 10-AUG-1995.

XX PF 10-JAN-1995; 95WO-US000296.

XX PR 03-FEB-1994; 94US-00191571.

XX PA (RICH) MERRELL DOW PHARM INC.

XX PI Owen TJ, Kudlacz EM, Buck SH, Harbeson SL;

XX DR WPI; 1995-336695/43.

XX CC New cyclic peptide derivs. - are neurokinin A and tachykinin antagonists
XX CC useful e.g. for treating asthma or as analgesics.

XX PS Claim 8; Page 69; 82pp; English.

XX CC The patent describes novel cyclic hexapeptide and octapeptide compounds
XX CC which are antagonists of neurokinin A and which are useful medically as
XX CC analgesics and for treating respiratory diseases such as asthma. The
XX CC present sequence represents a specifically preferred example of the new
XX CC peptides

XX SQ Sequence 6 AA;

Query Match 82.1%; Score 32; DB 2; Length 6;
Best Local Similarity 50.0%; Pred. No. 1.7e+06;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXWX 7
||::|:
Db 1 YWLRGW 6

RESULT 12
AAU83934
ID AAU83934 standard; peptide; 6 AA.

XX AC AAU83934;

XX XX

PD 23-DEC-1999.
XX
XX PF 15-JUN-1999; 99WO-IL000329.
XX
XX PR 19-JUN-1998; 98US-00100360.
PR 02-DEC-1998; 98US-00203389.
XX
XX PA (PEPT-) PEPTOR LTD.
XX
XX PI Hornik V, Afargan MM, Gellerman G;
XX WPI; 2000-136888/12.
DR
XX Cyclized somatostatin analogs for inhibiting growth hormone secretion
PT from anterior pituitary and as antiproliferative agents for the treatment
PT of tumors.
XX
XX Example 11; Page 61; 82pp; English.
XX
XX This sequence represents a somatostatin analogue of the invention. The
CC invention relates to a backbone cyclised somatostatin analogue that has
CC one building unit containing a nitrogen atom of the peptide backbone
CC connected to a bridging group comprising an amide, thioether, thioester
CC or disulphide. At least one building unit is connected via a bridging
CC group to form a cyclic structure with a moiety selected from a second
CC building unit, side chain of or N-terminal amino acid residue. A
CC composition containing the analogue may be used for preventing disorders
CC such as cancers, autoimmune diseases, endocrine disorders, diabetic-
CC associated complications, gastrointestinal disorders, inflammatory
CC diseases, pancreatitis, atherosclerosis, restenosis and post-surgical
CC pain. It may also be used for diagnosing cancer. The backbone cyclic
CC analogue is used for imaging the existence of metastases. Somatostatin
CC analogues can be used for the treatment of patients with hormone-secreting
CC and hormone-dependent tumours. They reduce diarrhoea through the
CC inhibition of vasoactive intestinal peptide (VIP) secretion and by direct
CC effect on intestinal secretion. Somatostatin analogues selective to type
CC 2 and 5 receptors may be used for treatment of non-insulin dependent
CC diabetes mellitus. They are useful for the prevention of atherosclerosis
CC and restenosis. The analogues are metabolically stable, selective in
CC their in-vivo activities and safe
XX
XX Sequence 7 AA;
Query Match 82.1%; Score 32; DB 3; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.7e+06;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 2 YWXXXW 7
Db 2 YWKVCW 7
RESULT 15
ABP53416
ID ABP53416 standard; peptide; 7 AA.
XX
XX AC ABP53416;
XX
XX DT 19-NOV-2002 (first entry)
XX
XX DE Backbone cyclised somatostatin analogue PTR 3177.
XX
XX KW Backbone cyclised somatostatin analogue; somatostatin; SRIF; analgesic;
KW somatostatin release inhibiting factor; somatostatin receptor subtype;
KW synthesis; antiatherosclerotic; immunosuppressive; cytostatic; cancer;
KW antidiabetic; antiinflammatory; somatostatin receptor ligand;
KW atherosclerosis; autoimmune disease; diabetic-associated complication;
KW endocrine disorder; inflammation; gastrointestinal disorder; restenosis;
KW pancreatitis; post-surgical pain.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers

FT Misc-difference 3 /note= "D form residue"
FT Modified-site 7
FT /note= "amidated"
XX
XX PN US2002052315-A1.
XX
XX PD 02-MAY-2002.
XX
XX PF 13-DEC-2000; 2000US-00734583.
XX
XX PR 19-JUN-1998; 98US-00100360.
PR 02-DEC-1998; 98US-00203389.
PR 15-JUN-1999; 99WO-IL000329.
XX
XX (HORN/) HORNIK V.
PA (AFAR/) AFARGAN M M.
PA (GELL/) GELLERMAN G.
XX
XX PI Hornik V, Afargan MM, Gellerman G;
XX WPI; 2002-681319/73.
XX
XX New backbone cyclized somatostatin analogs are e.g. useful in the
PT treatment of atherosclerosis, autoimmune diseases and cancers.
XX
XX Example 12; Page 21; 30pp; English.
XX
XX The present invention describes backbone cyclised somatostatin analogues
CC (I) that incorporates at least one building unit containing one nitrogen
CC atom of the peptide backbone connected to a bridging group (comprising an
CC amide, thioether, thioester or disulfide) where at least one building
CC unit is connected via the bridging group to form a cyclic structure with
CC a moiety selected from the group consisting of a second building unit,
CC the side chain of an amino acid residue of the sequence or the N-terminal
CC amino acid residue. (I) has antiatherosclerotic, immunosuppressive,
CC cytostatic, antidiabetic, antiinflammatory and analgesic activities, and
CC can be used as a somatostatin receptor ligand. (I) are useful in the
CC treatment of atherosclerosis, autoimmune diseases, cancers, diabetic-
CC associated complications, endocrine disorders, inflammation,
CC gastrointestinal disorders, pancreatitis, post-surgical pain, and
CC restenosis. (I) can also be used in the diagnosis of cancer, by imaging
CC the existence of metastases, it being labeled with a detectable probe.
CC The present sequence represents a backbone cyclised somatostatin analogue
CC from the present invention
XX
XX Sequence 7 AA;
Query Match 82.1%; Score 32; DB 5; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.7e+06;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 2 YWXXXW 7
Db 2 YWKVCW 7

Search completed: January 3, 2005, 17:32:20
Job time : 156 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 17:27:19 ; Search time 37 Seconds
(without alignments)
14.339 Million cell updates/sec

Title: US-10-046-922-68

Perfect score: 39

Sequence: 1 GYVXXWX 8

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 73656

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	84.6	7	4	US-09-069-827A-94
2	29	74.4	7	1	US-08-443-640-16
3	29	74.4	8	3	US-09-100-804-5
4	29	74.4	8	3	US-09-081-345-5
5	29	74.4	8	3	US-09-095-443-6
6	28	71.8	6	2	US-08-446-345-12
7	28	71.8	6	3	US-08-951-260A-2
8	28	71.8	6	3	US-08-884-569A-11
9	28	71.8	6	3	US-08-884-569A-14
10	28	71.8	6	4	US-09-430-626A-2
11	28	71.8	6	4	US-09-361-096A-46
12	28	71.8	6	4	US-09-743-492A-5
13	28	71.8	6	4	US-10-243-687-2
14	28	71.8	7	1	US-07-973-235A-30
15	28	71.8	7	2	US-08-652-971-5
16	28	71.8	7	2	US-08-462-720-30
17	28	71.8	7	2	US-08-991-258A-5
18	28	71.8	7	2	US-08-769-399-5
19	28	71.8	7	3	US-08-991-953A-5
20	28	71.8	8	3	US-08-467-472C-5
21	28	71.8	8	3	US-08-467-472C-6
22	28	71.8	8	3	US-08-467-472C-7
23	28	71.8	8	3	US-08-467-472C-8
24	28	71.8	8	3	US-09-384-061-5
25	28	71.8	8	3	US-09-384-061-6
26	28	71.8	8	3	US-09-384-061-7
27	28	71.8	8	3	US-09-384-061-8

28 71.8 8 4 US-09-794-529B-1 Sequence 1, Appli
29 71.8 8 4 US-09-794-517A-1 Sequence 1, Appli
30 71.8 8 4 US-09-011-645E-1 Sequence 1, Appli
31 71.8 8 4 US-09-794-832-1 Sequence 1, Appli
32 71.8 8 4 US-09-852-870A-5 Sequence 5, Appli
33 71.8 8 4 US-09-852-870A-7 Sequence 7, Appli
34 71.8 8 4 US-09-680-806A-1 Sequence 1, Appli
35 71.8 8 4 US-09-552-868-1 Sequence 1, Appli
36 71.8 8 4 US-09-636-295-1 Sequence 1, Appli
37 69.2 6 1 US-08-487-006-67 Sequence 67, Appl
38 69.2 6 2 US-08-488-659A-67 Sequence 67, Appl
39 69.2 7 4 US-09-563-222C-53 Sequence 53, Appl
40 69.2 8 3 US-08-925-002-42 Sequence 42, Appl
41 69.2 8 3 US-08-586-670A-17 Sequence 17, Appl
42 69.2 8 4 US-09-780-070-5 Sequence 5, Appli
43 69.2 8 4 US-09-910-552-42 Sequence 42, Appl
44 66.7 5 1 US-08-353-400-27 Sequence 27, Appl
45 66.7 6 1 US-07-718-577-6 Sequence 6, Appli

ALIGNMENTS

RESULT 1

US-09-069-827A-94
; Sequence 94, Application US/09069827A
; Patent No. 6617114
; GENERAL INFORMATION:
; APPLICANT: FOWLKES, Dana M
; KAY, Brian K
; FRELINGER, Jeffrey A
; HYDE-DERUYSCHE, Robin P
; TITLE OF INVENTION: IDENTIFICATION OF DRUGS USING
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEWMARK, P.L.L.C.
; STREET: 624 Ninth Street N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/069,827A
; FILING DATE: 30-Apr-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/050,359
; FILING DATE: 31-MAR-1998
; APPLICATION NUMBER: PCT/US97/19638
; FILING DATE: 31-OCT-1997
; APPLICATION NUMBER: US 08/740,671
; FILING DATE: 31-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Iver P
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: FOWLKES-4C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 94:

```
US-09-069-827A-94
Query Match      84.6%; Score 33; DB 4; Length 7;
Best Local Similarity 42.9%; Pred. No. 3.8e+05;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXWX 8
Db 1 YWPDWG 7

RESULT 2
US-08-443-640-16
; Sequence 16, Application US/08443640
; Patent No. 5691140
; GENERAL INFORMATION:
; APPLICANT: NOREN, CHRISTOPHER J.
; APPLICANT: EVANS, PAUL D.
; TITLE OF INVENTION: BIDIRECTIONAL IN VITRO TRANSCRIPTION
; TITLE OF INVENTION: VECTORS UTILIZING A SINGLE RNA POLYMERASE FOR BOTH
; DIRECTIONS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEW ENGLAND BIOLABS, INC
; STREET: 32 TOZER ROAD
; CITY: BEVERLY
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 01915
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/443,640
; FILING DATE: 18-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.
; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: NEB-102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 927-5054
; TELEFAX: (508) 927-1705
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-443-640-16

Query Match      74.4%; Score 29; DB 1; Length 7;
Best Local Similarity 42.9%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
Db 1 GWRPFW 7

RESULT 3
US-09-100-804-5
; Sequence 5, Application US/09100804
; Patent No. 6066472
; GENERAL INFORMATION:
; APPLICANT: GONEZ, LEONEL JORGE
; APPLICANT: SARAS, JAN
; APPLICANT: CLAESSON-WELSH, LENA
; APPLICANT: HELDIN, CARL-HENRIK
; TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL

US-09-069-827A-94
; TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,804
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/596,291
; FILING DATE: 09-AUG-1996
; APPLICATION NUMBER: US 08/115,573
; FILING DATE: 01-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/09943
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: L0461/7003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURES:
; NAME/KEY: Region
; LOCATION: 5
; OTHER INFORMATION: /note= "Xaa = I or V"
US-09-100-804-5

Query Match      74.4%; Score 29; DB 3; Length 8;
Best Local Similarity 42.9%; Pred. No. 3.8e+05;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXWX 8
Db 1 FWRMXWE 7

RESULT 4
US-09-081-345-5
; Sequence 5, Application US/09081345
; Patent No. 6228641
; GENERAL INFORMATION:
; APPLICANT: Bahija Jallal
; APPLICANT: Gregory D. Plowman
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: PTP04 RELATED DISORDERS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
```

;; CITY: Los Angeles
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 90071-2066
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;; MEDIUM TYPE: storage
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: IBM P.C. DOS 5.0
;; SOFTWARE: FastSeq for Windows 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/081,345
;; FILING DATE: Herewith
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/047,222
;; FILING DATE: May 20, 1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warburg, Richard J.
;; REGISTRATION NUMBER: 32,327
;; REFERENCE/DOCKET NUMBER: 234/253
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (213) 489-1600
;; TELEFAX: (213) 955-0440
;; TELEX: 67-3510
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 8 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FEATURE:
;; OTHER INFORMATION: "Xaa" in positions 4 and 6 stand
;; OTHER INFORMATION: for an unspecified amino acid.
;; OTHER INFORMATION: "Xaa" in position 8 stands for
;; OTHER INFORMATION: either Glu or Asp.
US-09-081-345-5

Query Match 74.4%; Score 29; DB 3; Length 8;
Best Local Similarity 71.4%; Pred. No. 3.8e+05;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXWX 8
Db 2 FWXXWX 8

RESULT 5
US-09-095-443-6
; Sequence 6, Application US/09095443
; Patent No. 6342593
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory
; APPLICANT: Peles, Eior
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF ALP RELATED DISORDERS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/09/095,443
;; FILING DATE: Herewith
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/049,477
;; FILING DATE: June 12, 1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warburg, Richard J.
;; REGISTRATION NUMBER: 32,327
;; REFERENCE/DOCKET NUMBER: 235/055
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (213) 489-1600
;; TELEFAX: (213) 955-0440
;; TELEX: 67-3510
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 8 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: Peptide
;; FEATURE:
;; OTHER INFORMATION: "Xaa" in positions 4 and 6 stand
;; OTHER INFORMATION: for an unspecified amino acid.
;; OTHER INFORMATION: "Xaa" in position 8 stands for
;; OTHER INFORMATION: either Glu or Asp.
US-09-095-443-6

Query Match 74.4%; Score 29; DB 3; Length 8;
Best Local Similarity 71.4%; Pred. No. 3.8e+05;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXWX 8
Db 2 FWXXWX 8

RESULT 6
US-08-446-345-12
; Sequence 12, Application US/08446345
; Patent No. 5831009
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; TITLE OF INVENTION: NOVEL PROTEIN PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASES PTP-D1
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: N.Y.
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,345
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/234,440
; FILING DATE: 28-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30742
; REFERENCE/DOCKET NUMBER: 7683-054
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090

```
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-446-345-12

Query Match 71.8%; Score 28; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
Db 1 FWXXW 6

RESULT 7
US-08-951-260A-2
; Sequence 2, Application US/08951260A
; Patent No. 6004791
; GENERAL INFORMATION:
; APPLICANT: Aoki, Naohito
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PROTEIN TYROSINE PHOSPHATASE PTP20
; TITLE OF INVENTION: AND RELATED PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951,260A
; FILING DATE: October 16, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,860
; FILING DATE: No. 6004791ember 13, 1996
; APPLICATION NUMBER: PCT/1897/00946
; FILING DATE: June 17, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: "Xaa" in positions 3 and 5 stand
; OTHER INFORMATION: for an unspecified amino acid.
US-08-951-260A-2

; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-446-345-12

Query Match 71.8%; Score 28; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
Db 1 FWXXW 6

RESULT 8
US-08-884-569A-11
; Sequence 11, Application US/08884569A
; Patent No. 6399326
; GENERAL INFORMATION:
; APPLICANT: CHIANG, MING-KO
; APPLICANT: FLANAGAN, JOHN G.
; TITLE OF INVENTION: RECEPTOR TYROSINE PHOSPHATASE, AND USES RELATED THERETO
; FILE REFERENCE: HMV-020.01
; CURRENT APPLICATION NUMBER: US/08/884,569A
; CURRENT FILING DATE: 1997-06-27
; PRIOR APPLICATION NUMBER: 60/021,040
; PRIOR FILING DATE: 1996-07-02
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: MOD_RES
; LOCATION: (5)
; OTHER INFORMATION: Ile or Val
US-08-884-569A-11

Query Match 71.8%; Score 28; DB 3; Length 6;
Best Local Similarity 50.0%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
Db 1 FWXXW 6

RESULT 9
US-08-884-569A-14
; Sequence 14, Application US/08884569A
; Patent No. 6399326
; GENERAL INFORMATION:
; APPLICANT: CHIANG, MING-KO
; APPLICANT: FLANAGAN, JOHN G.
; TITLE OF INVENTION: RECEPTOR TYROSINE PHOSPHATASE, AND USES RELATED THERETO
; FILE REFERENCE: HMV-020.01
; CURRENT APPLICATION NUMBER: US/08/884,569A
; CURRENT FILING DATE: 1997-06-27
; PRIOR APPLICATION NUMBER: 60/021,040
; PRIOR FILING DATE: 1996-07-02
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: MOD_RES
; LOCATION: (3)
; OTHER INFORMATION: Arg or Gln
; NAME/KEY: MOD_RES
; LOCATION: (5)
; OTHER INFORMATION: Ile or Val
```


US-08-884-569A-14

Query Match 71.8%; Score 28; DB 3; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
:|:|:
Db 1 FWXXW 6

RESULT 10

US-09-430-626A-2
; Sequence 2, Application US/09430626A
; Patent No. 6482605
; GENERAL INFORMATION:
; APPLICANT: Aoki, Naohito
; Ullrich, Axel
; TITLE OF INVENTION: PROTEIN TYROSINE PHOSPHATASE PTP20
; AND RELATED PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,626A
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/951,260
; FILING DATE: October 16, 1997
; APPLICATION NUMBER: 60/030,860
; FILING DATE: No. 6482605ember 13, 1996
; APPLICATION NUMBER: PCT/1897/00946
; FILING DATE: June 17, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

SEQUENCE CHARACTERISTICS:
INFORMATION FOR SEQ ID NO: 2:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: "Xaa" in positions 3 and 5 stand
for an unspecified amino acid.
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-430-626A-2

Query Match 71.8%; Score 28; DB 4; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
:|:|:
Db 1 FWXXW 6

Db

1 FWXXW 6

RESULT 11

US-09-361-096A-46
; Sequence 46, Application US/09361096A
; Patent No. 6492495
; GENERAL INFORMATION:
; APPLICANT: MOLLER, NIELS P.H.
; APPLICANT: MOLLER, KARIN B.
; APPLICANT: ULLRICH, AXEL
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE PHOSPHATASE
; FILE REFERENCE: 038602/0686
; CURRENT APPLICATION NUMBER: US/09/361,096A
; CURRENT FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 08/449,609
; PRIOR FILING DATE: 1995-05-24
; PRIOR APPLICATION NUMBER: 08/036,210
; PRIOR FILING DATE: 1995-03-23
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Consensus
; OTHER INFORMATION: sequence
; NAME/KEY: MOD_RES
; LOCATION: (3)
; OTHER INFORMATION: Variable amino acid
; NAME/KEY: MOD_RES
; LOCATION: (5)
; OTHER INFORMATION: Variable amino acid
; US-09-361-096A-46

Query Match 71.8%; Score 28; DB 4; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
:|:|:
Db 1 FWXXW 6

RESULT 12

US-09-743-492A-5
; Sequence 5, Application US/09743492A
; Patent No. 6709843
; GENERAL INFORMATION:
; APPLICANT: YAMAMOTO, Hiroshi
; APPLICANT: TSUJIKAWA, Kazutake
; APPLICANT: UCHINO, Yukiko
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR INTRACELLULAR DOMAIN OF PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASE
; FILE REFERENCE: 19036/37023
; CURRENT APPLICATION NUMBER: US/09/743,492A
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: PCT/JP98/03120
; PRIOR FILING DATE: 1998-07-10
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Consensus Amino
; OTHER INFORMATION: Acid Sequence in Cytoplasmic Domain of Known PTPs.
; Patent No. 6709843
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)

OTHER INFORMATION: "Xaa" in positions 3 and 5 stand
for an unspecified amino acid.
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-430-626A-2

; OTHER INFORMATION: Xaa= Arg, Glu or Leu
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa= Val, Ile or Cys
US-09-743-492A-5

Query Match 71.8%; Score 28; DB 4; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7
:|:|:
Db 1 FWXXW 6

RESULT 13
US-10-243-687-2
; Sequence 2, Application US/10243687
; Patent No. 6797501
; GENERAL INFORMATION:
; APPLICANT: Aoki, Naohito
; Ullrich, Axel
; TITLE OF INVENTION: PROTEIN TYROSINE PHOSPHATASE PTP20
; AND RELATED PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/243,687
FILING DATE: 16-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/430,626A
FILING DATE: 29-Oct-1999
APPLICATION NUMBER: 08/951,260
FILING DATE: October 16, 1997
APPLICATION NUMBER: 60/030,860
FILING DATE: No. 6797501ember 13, 1996
APPLICATION NUMBER: PCT/1897/00946
FILING DATE: June 17, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: "Xaa" in positions 3 and 5 stand
for an unspecified amino acid.
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-243-687-2

Query Match 71.8%; Score 28; DB 4; Length 6;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7
:|:|:
Db 1 FWXXW 6

RESULT 14
US-07-973-235A-30
; Sequence 30, Application US/07973235A
; Patent No. 5491130
; GENERAL INFORMATION:
; APPLICANT: David D. Roberts, et al.
; TITLE OF INVENTION: Peptide Inhibitors of Fibronectin and
; TITLE OF INVENTION: Related Collagen-Binding Proteins
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Lowe, Price, LeBlanc & Becker
; STREET: Suite 300, 99 Canal Center Plaza
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect, Version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/973,235A
FILING DATE: 19921110
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Robert L. Price
REGISTRATION NUMBER: 22,685
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)684-1111
TELEFAX: (703)684-1124
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: AMINO ACIDS
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-973-235A-30

Query Match 71.8%; Score 28; DB 1; Length 7;
Best Local Similarity 42.9%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYVXXXW 7
|:|:|:
Db 1 GGWSKSW 7

RESULT 15
US-08-652-971-5
; Sequence 5, Application US/08652971
; Patent No. 5814507
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; APPLICANT: Lasky, Laurence A.
; TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE, PTP LAMEDA
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd.
; CITY: South San Francisco

STATE: California
COUNTRY: United States
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,971
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P1033
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 225-3216
TELEFAX: (415) 952-9881
TELEX: 910 371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Active-site
LOCATION: 1..2
OTHER INFORMATION: /note= "Let 'X' located at position
OTHER INFORMATION: 1 represent either Histidine or Aspartic Acid"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6..7
OTHER INFORMATION: /note= "Let 'X' located at position
OTHER INFORMATION: 6 represent either Isoleucine or Valine."
US-08-652-971-5

Query Match 71.8%; Score 28; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 3.8e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXW 7
Db 2 FWRXW 7

Search completed: January 3, 2005, 17:36:59
Job time : 38 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 17:35:40 ; Search time 140 Seconds
(without alignments)
20.556 Million cell updates/sec

Title: US-10-046-922-68
Perfect score: 39
Sequence: 1 GYXXXXXX 8

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 83711

Minimum DB seq length: 0
Maximum-DB-seq-length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	8	13	US-10-046-922-68
2	38	97.4	7	13	Sequence 68, Appl
3	32	82.1	6	15	US-10-046-922-67
4	32	82.1	8	14	US-10-418-943-48
5	30	76.9	8	15	US-10-190-082-586
6	30	76.9	8	15	Sequence 586, App
7	30	76.9	8	15	US-10-367-580-243
8	30	76.9	8	15	Sequence 243, App
9	30	76.9	8	15	US-10-367-593-243
10	30	76.9	8	15	Sequence 243, App
11	30	76.9	8	15	US-10-367-654-243
12	29	74.4	8	16	US-10-367-658-243
13	29	74.4	8	10	US-10-367-674-243
					Sequence 243, App
					Sequence 5, Appli
					Sequence 12, Appl

14	74.4	8	14	US-10-020-215-6	Sequence 6, Appli
15	74.4	8	15	US-10-328-953-9	Sequence 9, Appli
16	74.4	8	15	US-10-367-580-153	Sequence 153, App
17	74.4	8	15	US-10-367-593-153	Sequence 153, App
18	74.4	8	15	US-10-367-594-153	Sequence 153, App
19	74.4	8	15	US-10-367-654-153	Sequence 153, App
20	74.4	8	15	US-10-367-658-153	Sequence 153, App
21	74.4	8	15	US-10-367-668-153	Sequence 153, App
22	74.4	8	16	US-10-367-674-153	Sequence 153, App
23	71.8	6	13	US-10-087-993-1	Sequence 1, Appli
24	71.8	6	14	US-10-243-687-2	Sequence 2, Appli
25	71.8	6	14	US-10-314-232-46	Sequence 46, Appl
26	71.8	7	14	US-10-190-082-65	Sequence 65, Appl
27	71.8	7	14	US-10-190-082-76	Sequence 76, Appl
28	71.8	7	17	US-10-858-271-10	Sequence 10, Appl
29	71.8	8	9	US-09-852-870A-5	Sequence 5, Appli
30	71.8	8	9	US-09-852-870A-7	Sequence 7, Appli
31	71.8	8	10	US-09-794-517-1	Sequence 1, Appli
32	71.8	8	10	US-09-794-529-1	Sequence 1, Appli
33	71.8	8	10	US-09-794-832-1	Sequence 1, Appli
34	71.8	8	14	US-10-052-578-143	Sequence 143, App
35	71.8	8	14	US-10-053-520-143	Sequence 143, App
36	71.8	8	14	US-10-170-713A-1	Sequence 1, Appli
37	71.8	8	14	US-10-171-734-1	Sequence 1, Appli
38	71.8	8	14	US-10-053-498B-143	Sequence 143, App
39	71.8	8	15	US-10-258-147-3	Sequence 3, Appli
40	71.8	8	15	US-10-258-146A-1	Sequence 1, Appli
41	71.8	8	15	US-10-328-953-3	Sequence 3, Appli
42	71.8	8	15	US-10-367-580-1	Sequence 1, Appli
43	71.8	8	15	US-10-367-593-1	Sequence 1, Appli
44	71.8	8	15	US-10-367-594-1	Sequence 1, Appli
45	71.8	8	15	US-10-367-654-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-10-046-922-68
; Sequence 68, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 68
; LENGTH: 8
; TYPE: PRT
; ORGANISM: peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)..(6)
; OTHER INFORMATION: X is any amino acid
; NAME/KEY: SITE
; LOCATION: (8)..(8)
; OTHER INFORMATION: X is any amino acid
US-10-046-922-68

Query Match 100.0%; Score 39; DB 13; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8
| | | | | | | |
Db 1 GYXXXXXX 8

RESULT 2
US-10-046-922-67
; Sequence 67, Application US/10046922
; Publication No. US20020164667A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Koivunen, Erkki
; APPLICANT: Kubo, Hajime
; TITLE OF INVENTION: VEGFR-3 INHIBITOR MATERIALS AND METHODS
; FILE REFERENCE: 28967/37084A
; CURRENT APPLICATION NUMBER: US/10/046,922
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 67
; LENGTH: 7
; TYPE: PRT
; ORGANISM: peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)..(6)
; OTHER INFORMATION: X at position 4-6 is any amino acid
US-10-046-922-67

Query Match 97.4%; Score 38; DB 13; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
| | | | |
Db 1 GYWXXXW 7

RESULT 3
US-10-418-943-48
; Sequence 48, Application US/10418943
; Publication No. US2004000241A1
; GENERAL INFORMATION:
; APPLICANT: Segall, Anca
; APPLICANT: Finilla, Clemencia
; TITLE OF INVENTION: RECOMBINATION MODULATORS AND METHODS
; TITLE OF INVENTION: FOR PRODUCING AND USING THE SAME
; FILE REFERENCE: 011443 008-999
; CURRENT APPLICATION NUMBER: US/10/418,943
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: 09/602,087
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-418-943-48

Query Match 82.1%; Score 32; DB 15; Length 6;
Best Local Similarity 50.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7
| | | | |
Db 1 YWCYWW 6

RESULT 4
US-10-190-082-586
; Sequence 586, Application US/10190082
; Publication No. US20030148264A1
; GENERAL INFORMATION:
; APPLICANT: Lasky, Lawrence A.
; APPLICANT: Sidhu, Sachdev S.
; APPLICANT: Held, Heike A.
; TITLE OF INVENTION: PHAGE DISPLAYED PDZ DOMAIN LIGANDS

FILE REFERENCE: P1905R1
; CURRENT APPLICATION NUMBER: US/10/190,082
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,634
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 683
; SEQ ID NO 586
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-190-082-586

Query Match 82.1%; Score 32; DB 14; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXXW 7
| | | | |
Db 3 YWEYWW 8

RESULT 5
US-10-367-580-243
; Sequence 243, Application US/10367580
; Publication No. US20040071720A1
; GENERAL INFORMATION:
; APPLICANT: Rothman, James E.
; APPLICANT: Hartl, F. Ulrich
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Takechi, Yoshizumi
; APPLICANT: Mayhew, Mark
; TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
; FILE REFERENCE: 11746/461061
; CURRENT APPLICATION NUMBER: US/10/367,580
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 09/794,832
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: US 09/011,645
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: PCT/US96/13363
; PRIOR FILING DATE: 1996-08-16
; PRIOR APPLICATION NUMBER: US 60/002,490
; PRIOR FILING DATE: 1995-08-18
; PRIOR APPLICATION NUMBER: US 60/002,479
; PRIOR FILING DATE: 1995-08-18
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 243
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-367-580-243

Query Match 76.9%; Score 30; DB 15; Length 8;
Best Local Similarity 42.9%; Pred. No. 1.5e+06;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWXXXW 7
| | | | |
Db 2 GLWFFPW 8

RESULT 6
US-10-367-593-243
; Sequence 243, Application US/10367593
; Publication No. US20040071721A1
; GENERAL INFORMATION:
; APPLICANT: Rothman, James E.

APPLICANT: Hartl, F. Ulrich
APPLICANT: Hoe, Mee H.
APPLICANT: Houghton, Alan
APPLICANT: Takechi, Yoshizumi
APPLICANT: Mayhew, Mark
TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
FILE REFERENCE: 11746/461012
CURRENT APPLICATION NUMBER: US/10/367,593
CURRENT FILING DATE: 2003-02-14
PRIOR APPLICATION NUMBER: US 09/011,645
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: PCT/US96/13363
PRIOR FILING DATE: 1996-08-16
PRIOR APPLICATION NUMBER: US 60/002,490
PRIOR FILING DATE: 1995-08-18
PRIOR APPLICATION NUMBER: US 60/002,479
PRIOR FILING DATE: 1995-08-18
NUMBER OF SEQ ID NOS: 349
SOFTWARE: WordPerfect 8.0 for Windows
SEQ ID NO 243
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-10-367-593-243

Query Match 76.9%; Score 30; DB 15; Length 8;
Best Local Similarity 42.9%; Pred. No. 1.5e+06;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYWXW 7
| |::|
Db 2 GLWFFW 8

RESULT 7
US-10-367-594-243
Sequence 243, Application US/10367594
Publication No. US2004007172A1
GENERAL INFORMATION:
APPLICANT: Rothman, James E.
APPLICANT: Hartl, F. Ulrich
APPLICANT: Hoe, Mee H.
APPLICANT: Houghton, Alan
APPLICANT: Takechi, Yoshizumi
APPLICANT: Mayhew, Mark
TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
FILE REFERENCE: 11746/461041
CURRENT APPLICATION NUMBER: US/10/367,594
CURRENT FILING DATE: 2003-02-14
PRIOR APPLICATION NUMBER: US 09/680,806
PRIOR FILING DATE: 2000-10-05
PRIOR APPLICATION NUMBER: US 09/011,645
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: PCT/US96/13363
PRIOR FILING DATE: 1996-08-16
PRIOR APPLICATION NUMBER: US 60/002,490
PRIOR FILING DATE: 1995-08-18
PRIOR APPLICATION NUMBER: US 60/002,479
PRIOR FILING DATE: 1995-08-18
NUMBER OF SEQ ID NOS: 349
SOFTWARE: WordPerfect 8.0 for Windows
SEQ ID NO 243
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-10-367-594-243

Query Match 76.9%; Score 30; DB 15; Length 8;
Best Local Similarity 42.9%; Pred. No. 1.5e+06;

Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GYWXW 7
| |::|
Db 2 GLWFFW 8

RESULT 8
US-10-367-654-243
Sequence 243, Application US/10367654
Publication No. US2004007172A1
GENERAL INFORMATION:
APPLICANT: Rothman, James E.
APPLICANT: Hartl, F. Ulrich
APPLICANT: Hoe, Mee H.
APPLICANT: Houghton, Alan
APPLICANT: Takechi, Yoshizumi
APPLICANT: Mayhew, Mark
TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
FILE REFERENCE: 11746/461032
CURRENT APPLICATION NUMBER: US/10/367,654
CURRENT FILING DATE: 2003-02-14
PRIOR APPLICATION NUMBER: US 10/171,734
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: US 09/636,295
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US 09/011,645
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: PCT/US96/13363
PRIOR FILING DATE: 1996-08-16
PRIOR APPLICATION NUMBER: US 60/002,490
PRIOR FILING DATE: 1995-08-18
PRIOR APPLICATION NUMBER: US 60/002,479
PRIOR FILING DATE: 1995-08-18
NUMBER OF SEQ ID NOS: 349
SOFTWARE: WordPerfect 8.0 for Windows
SEQ ID NO 243
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-10-367-654-243

Query Match 76.9%; Score 30; DB 15; Length 8;
Best Local Similarity 42.9%; Pred. No. 1.5e+06;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYWXW 7
| |::|
Db 2 GLWFFW 8

RESULT 9
US-10-367-658-243
Sequence 243, Application US/10367658
Publication No. US2004007172A1
GENERAL INFORMATION:
APPLICANT: Rothman, James E.
APPLICANT: Hartl, F. Ulrich
APPLICANT: Hoe, Mee H.
APPLICANT: Houghton, Alan
APPLICANT: Takechi, Yoshizumi
APPLICANT: Mayhew, Mark
TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
FILE REFERENCE: 11746/461051
CURRENT APPLICATION NUMBER: US/10/367,658
CURRENT FILING DATE: 2003-02-14
PRIOR APPLICATION NUMBER: US 09/794,529
PRIOR FILING DATE: 2001-02-27
PRIOR APPLICATION NUMBER: US 09/011,645
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: PCT/US96/13363

; PRIOR FILING DATE: 1996-08-16
; PRIOR APPLICATION NUMBER: US 60/002,490
; PRIOR FILING DATE: 1995-08-18
; PRIOR APPLICATION NUMBER: US 60/002,479
; PRIOR FILING DATE: 1995-08-18
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 243
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-367-658-243

Query Match 76.9%; Score 30; DB 15; Length 8;
Best Local Similarity 42.9%; Pred. No. 1.5e+06;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWXXW 7
| |::|
Db 2 GLWFFW 8

RESULT 10
US-10-367-668-243
; Sequence 243, Application US/10367668
; Publication No. US20040071725A1
; GENERAL INFORMATION:
; APPLICANT: Rothman, James E.
; APPLICANT: Hartl, F. Ulrich
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Takechi, Yoshizumi
; APPLICANT: Mayhew, Mark
; TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
; FILE REFERENCE: 11746/461072
; CURRENT APPLICATION NUMBER: US/10/367,668
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 09/794,517
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: US 09/011,645
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: PCT/US96/13363
; PRIOR FILING DATE: 1996-08-16
; PRIOR APPLICATION NUMBER: US 60/002,490
; PRIOR FILING DATE: 1995-08-18
; PRIOR APPLICATION NUMBER: US 60/002,479
; PRIOR FILING DATE: 1995-08-18
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 243
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-367-668-243

Query Match 76.9%; Score 30; DB 15; Length 8;
Best Local Similarity 42.9%; Pred. No. 1.5e+06;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWXXW 7
| |::|
Db 2 GLWFFW 8

RESULT 11
US-10-367-674-243
; Sequence 243, Application US/10367674
; Publication No. US20040127684A1
; GENERAL INFORMATION:

; APPLICANT: Rothman, James E.
; APPLICANT: Hartl, F. Ulrich
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Takechi, Yoshizumi
; APPLICANT: Mayhew, Mark
; TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
; FILE REFERENCE: 11746/4610211
; CURRENT APPLICATION NUMBER: US/10/367,674
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 10/170,738
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/552,868
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 09/011,645
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: PCT/US96/13363
; PRIOR FILING DATE: 1996-08-16
; PRIOR APPLICATION NUMBER: US 60/002,490
; PRIOR FILING DATE: 1995-08-18
; PRIOR APPLICATION NUMBER: US 60/002,479
; PRIOR FILING DATE: 1995-08-18
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 243
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-367-674-243

Query Match 76.9%; Score 30; DB 16; Length 8;
Best Local Similarity 42.9%; Pred. No. 1.5e+06;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWXXW 7
| |::|
Db 2 GLWFFW 8

RESULT 12
US-09-822-295-5
; Sequence 5, Application US/09822295
; Patent No. US20020119501A1
; GENERAL INFORMATION:
; APPLICANT: Bahija Jallal
; Gregory D. Plowman
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
PTP04 RELATED DISORDERS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/822,295
; FILING DATE: 02-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/081,345
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 234/253
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: "Xaa" in positions 4 and 6 stand
for an unspecified amino acid.
"Xaa" in position 8 stands for
either Glu or Asp.
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-822-295-5

Query Match 74.4%; Score 29; DB 9; Length 8;
Best Local Similarity 71.4%; Pred. No. 1.5e+06;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXYWX 8
:|:|:|
Db 2 FWXXYWX 8

RESULT 13

US-09-095-478-12
; Sequence 12, Application US/09095478
; Publication No. US20030095970A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, Gregory
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE SUPT905 AND
; TITLE OF INVENTION: RELATED PRODUCTS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,478
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 224/115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
FEATURE:
OTHER INFORMATION: "Xaa" in positions 4 and 6 stand
for an unspecified amino acid.
"Xaa" in position 8 stands for
either Glu or Asp.
US-09-095-478-12

Query Match 74.4%; Score 29; DB 10; Length 8;
Best Local Similarity 71.4%; Pred. No. 1.5e+06;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXYWX 8
:|:|:|
Db 2 FWXXYWX 8

RESULT 14

US-10-020-215-6
; Sequence 6, Application US/10020215
; Publication No. US2003008347A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: PELES, EIOR
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF ALP RELATED DISORDERS
; FILE REFERENCE: 038602/1290
; CURRENT APPLICATION NUMBER: US/10/020,215
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 09/095,443
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/049,477
; PRIOR FILING DATE: 1997-06-11
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (4)
; OTHER INFORMATION: Unspecified amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (6)
; OTHER INFORMATION: Unspecified amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (8)
; OTHER INFORMATION: Glu or Asp
US-10-020-215-6
Query Match 74.4%; Score 29; DB 14; Length 8;
Best Local Similarity 71.4%; Pred. No. 1.5e+06;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXYWX 8
:|:|:|
Db 2 FWXXYWX 8

RESULT 15
US-10-328-953-9
; Sequence 9, Application US/10328953
; Publication No. US20040071656A1
; GENERAL INFORMATION:

```

; APPLICANT: Wieland, Felix
; APPLICANT: Hartl, Franz-Ulrich
; TITLE OF INVENTION: Modulation of Heat-Shock-Protein-Based Immunotherapies
; FILE REFERENCE: 11390/46101
; CURRENT APPLICATION NUMBER: US/10/328,953
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 60/342,570
; PRIOR FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: US 60/343,884
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/372,620
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US 60/399,342
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: US 60/414,834
; PRIOR FILING DATE: 2002-09-28
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 9
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide in M13 coliphage
US-10-328-953-9

```

```

Query Match      74.4%; Score 29; DB 15; Length 8;
Best Local Similarity 28.6%; Pred. No. 1.5e+06;
Matches 2; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 YWXXWX 8
Db      2 FWPFEWI 8

```

```

Search completed: January 3, 2005, 17:47:30
Job time : 141 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 17:26:39 ; Search time 37 seconds
(without alignments)
20.804 Million cell updates/sec

Title: US-10-046-922-68

Perfect score: 39

Sequence: 1 GYWXKWX 8

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 606

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	56.4	5	JH0253	gut pentapeptide -
2	17	43.6	6	B35640	cerebellar degener
3	17	43.6	6	B35640	T-cell receptor ga
4	17	43.6	7	S21230	dermorphin (trp-4,
5	17	43.6	8	JS0315	leucokinin V - Mad
6	17	43.6	8	T13818	cytochrome oxidase
7	16	41.0	7	PT0728	T-cell receptor be
8	16	41.0	7	PN0649	pullulanase (EC 3.
9	16	41.0	8	S15422	adipokinetic hormo
10	16	41.0	8	A58641	adipokinetic hormo
11	16	41.0	8	S21663	neuropeptide - flo
12	15	38.5	5	PT0281	IG heavy chain CRD
13	15	38.5	5	PT0580	T-cell receptor be
14	15	38.5	6	S68195	alcohol dehydrogen
15	15	38.5	6	PT0629	T-cell receptor be
16	15	38.5	6	PT0532	T-cell receptor be
17	15	38.5	6	PT0519	T-cell receptor be
18	15	38.5	6	PT0637	T-cell receptor be
19	15	38.5	6	PT0641	T-cell receptor be
20	15	38.5	6	PD0028	pev-kinin 2 - pena
21	15	38.5	6	A61068	locustakinin - mig
22	15	38.5	7	S09652	hypothetical prote
23	15	38.5	7	PH1602	IG H chain V-D-J r
24	15	38.5	7	PT0526	T-cell receptor be
25	15	38.5	7	PT0628	T-cell receptor be
26	15	38.5	7	PT0642	T-cell receptor be
27	15	38.5	7	PT0722	T-cell receptor be
28	15	38.5	7	PT0586	T-cell receptor be
29	15	38.5	7	PX0008	glucuronosyltransf

ALIGNMENTS

RESULT 1

JH0253

gut pentapeptide - Japanese eel

C:Species: Anguilla japonica (Japanese eel)

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Apr-1995

C:Accession: JH0253

R:Uesaka, T.; Ikeda, T.; Kubota, I.; Muneoka, Y.; Ando, M.

Biochem. Biophys. Res. Commun. 180, 828-832, 1991

A:Title: Structure and function of a pentapeptide isolated from the gut of the eel.

A:Reference number: JH0253; MUID:92062113; PMID:1953755

A:Accession: JH0253

A:Molecule type: protein

A:Residues: 1-5 <UES>

A:Experimental source: gut

C:Comment: This peptide increased basal tone of the circular muscle of the esophagogastric

, and of the circular muscle of the gastro-intestinal junction.

Query Match 56.4%; Score 22; DB 2; Length 5;

Best Local Similarity 40.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWX 5

DB 1 GFWNK 5

RESULT 2

B35640

cerebellar degeneration-related protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 24-Jun-1993

C:Accession: B35640

R:Chen, Y.T.; Rettig, W.J.; Yenamandra, A.K.; Kozak, C.A.; Chaganti, R.S.K.; Posner, J.B.

Proc. Natl. Acad. Sci. U.S.A. 87, 3077-3081, 1990

A:Title: Cerebellar degeneration-related antigen: a highly conserved neuroectodermal mar

A:Reference number: A35640; MUID:90222173; PMID:2326268

A:Accession: B35640

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-6 <CHE>

Query Match

Best Local Similarity 43.6%; Score 17; DB 2; Length 6;

Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXX 6

DB 1 FWEDL 5

RESULT 3

F41946

T-cell receptor gamma chain (1a.27) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
 C;Accession: F41946
 R;WhetSELL, M.; Mosley, R.L.; WhetSELL, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
 Mol. Cell. Biol. 11, 5902-5909, 1991
 A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma gene
 A;Reference number: A41946; MUID:92049316; PMID:1658619
 A;Accession: F41946
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-6 <WHE>
 C;Keywords: T-cell receptor

Query Match 43.6%; Score 17; DB 2; Length 6;
 Best Local Similarity 16.7%; Pred. No. 2.8e+05;
 Matches 1; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXWX 8
 :|::|:
 Db 1 YCAVWV 6

RESULT 4
 S21230
 dermorphin (Trp-4, Asn-7) [validated] - two-colored leaf frog (fragment)
 C;Species: Phyllomedusa bicolor (two-colored leaf frog)
 C;Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 18-Aug-2000
 C;Accession: S21230
 R;Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Kreil, G.
 PEST Lett. 302, 151-154, 1992
 A;Title: Identification and characterization of two dermorphins from skin extracts of the
 A;Reference number: S21152; MUID:92339502; PMID:1633846
 A;Accession: S21230
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-7 <MIG>
 C;Superfamily: dermorphin precursor; dermorphin precursor amino-terminal homology

Query Match 43.6%; Score 17; DB 2; Length 7;
 Best Local Similarity 20.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWXXX 6
 :|::|:
 Db 3 FWYPN 7

RESULT 5
 JS0315
 leucokinin V - Madeira cockroach
 C;Species: Leucophaea maderae (Madeira cockroach)
 C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
 C;Accession: JS0315
 R;Holman, G.M.; Cook, B.J.; Nachman, R.J.
 Comp. Biochem. Physiol. C 88, 27-30, 1987
 A;Title: Isolation, primary structure, and synthesis of leucokinins V and VI: myotropic
 A;Reference number: JS0315
 A;Accession: JS0315
 A;Molecule type: protein
 A;Residues: 1-8 <HOL>
 A;Cross-references: UNIPROT:P19987
 C;Comment: Leucokinins, a family of cephalomyotropic peptides, stimulate contractile act
 C;Keywords: amidated carboxyl end; cephalomyotropic peptide
 F;8/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 43.6%; Score 17; DB 2; Length 8;
 Best Local Similarity 25.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWXXXWX 8
 :|::|:
 Db 1 GSGFSSWG 8

RESULT 6
 T13818
 cytochrome oxidase subunit I - Atlantic hagfish mitochondrion (fragment)
 C;Species: mitochondrion Myxine glutinosa (Atlantic hagfish)
 C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C;Accession: T13818
 R;Delarbre, C.; Barriel, V.; Tillier, S.; Janvier, P.; Gachelin, G.
 Mol. Biol. Evol. 14, 807-813, 1997
 A;Title: The main features of the craniate mitochondrial DNA between the ND1 and the COI
 A;Reference number: Z17775; MUID:97398704; PMID:9254518
 A;Accession: T13818
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-8
 A;Cross-references: UNIPROT:O21079; EMBL:Y09527; NID:g2340019; PIDN:CAA70718.1; PID:g2340
 C;Genetics:
 A;Genome: mitochondrion
 A;Note: COI
 C;Keywords: mitochondrion

Query Match 43.6%; Score 17; DB 2; Length 8;
 Best Local Similarity 16.7%; Pred. No. 2.8e+05;
 Matches 1; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXWX 8
 :|::|:
 Db 2 YLSRWF 7

RESULT 7
 PT0728
 T-cell receptor beta chain V-D-J region (161-2H) - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C;Accession: PT0728
 R;Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A;Reference number: PT0509; MUID:91277601; PMID:1711558
 A;Accession: PT0728
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-7 <FEE>
 A;Experimental source: newborn thymus, strain BALB/c
 C;Keywords: T-cell receptor

Query Match 41.0%; Score 16; DB 2; Length 7;
 Best Local Similarity 40.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWX 5
 :|::|:
 Db 3 GDMGG 7

RESULT 8
 PN0649
 pullulanase (EC 3.2.1.41) - Bacillus sp. (strain S-1) (fragment)
 C;Species: Bacillus sp.
 C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 03-Jun-2002
 C;Accession: PN0649
 R;Kim, C.H.; Choi, H.I.; Lee, D.S.
 Biosci. Biotechnol. Biochem. 57, 1632-1637, 1993
 A;Title: Purification and biochemical properties of an alkaline pullulanase from alkalop
 A;Reference number: PN0649; MUID:94080025; PMID:7764261
 A;Accession: PN0649
 A;Molecule type: protein
 A;Residues: 1-7 <KIM>
 C;Comment: This enzyme is used together with glucoamylase to improve the efficiency of sta
 nent in high maltose syrups.
 C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 41.0%; Score 16; DB 2; Length 7;
Best Local Similarity 14.3%; Pred. No. 2.8e+05;
Matches 1; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 YWXXWX 8
DB 1 FLNMSWF 7

RESULT 9
S15422
A:Title: adipokinetic hormone - cockchafer
C:Species: Melolontha melolontha (cockchafer)
C>Date: 19-Mar-1997 #sequence_revision 31-Oct-1997 #text_change 09-Jul-2004
C:Accession: S15422
R:Gaede, G.; 275, 671-677, 1991

Biochem. J. 275, 671-677, 1991
A:Title: A unique charged tyrosine-containing member of the adipokinetic hormone/ red-pi
A:Reference number: S15422; MUID:91248100; PMID:2039445
A:Accession: S15422
A:Molecule type: protein
A:Residues: 1-8 <BIO>

A:Cross-references: UNIPROT:P25423
C:Superfamily: adipokinetic hormone
C:Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 41.0%; Score 16; DB 2; Length 8;
Best Local Similarity 20.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXWX 7
DB 4 YSPDW 8

RESULT 10
A58641
A:Title: adipokinetic hormone - dor beetle
C:Species: Geotrupes stercorarius (dor beetle)
C>Date: 28-Oct-1997 #sequence_revision 31-Oct-1997 #text_change 09-Jul-2004
C:Accession: A58641
R:Gaede, G.

Biochem. J. 275, 671-677, 1991
A:Title: A unique charged tyrosine-containing member of the adipokinetic hormone/ red-pi
A:Reference number: S15422; MUID:91248100; PMID:2039445
A:Accession: A58641
A:Molecule type: protein
A:Residues: 1-8 <BIO>

A:Cross-references: UNIPROT:P25423
C:Superfamily: adipokinetic hormone
C:Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 41.0%; Score 16; DB 2; Length 8;
Best Local Similarity 20.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXWX 7
DB 4 YSPDW 8

RESULT 11
S21663
A:Title: neuropeptide - flower beetle (Pachnoda marginata)
C:Species: Pachnoda marginata
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S21663
R:Gaede, G.; Lopata, A.; Kellner, R.; Rinehart, K.L.

Biol. Chem. Hoppe-Seyler 373, 133-142, 1992
A:Title: Primary structures of neuropeptides isolated from the corpora cardiaca of various
ectometry.

A:Reference number: S21663; MUID:92265187; PMID:1586453
A:Accession: S21663
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <GAE>
A:Cross-references: UNIPROT:P25423

Query Match 41.0%; Score 16; DB 2; Length 8;
Best Local Similarity 20.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXWX 7
DB 4 YSPDW 8

RESULT 12
PT0281
A:Title: Ig heavy chain CRD3 region (clone 4-91C) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0281
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 195-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j

A:Reference number: PT0222; MUID:91108337; PMID:1899102
A:Accession: PT0281
A:Molecule type: DNA
A:Residues: 1-5 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match 38.5%; Score 15; DB 2; Length 5;
Best Local Similarity 20.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 4 XXXWX 8
DB 1 DENWS 5

RESULT 13
PT0580
A:Title: T-cell receptor beta chain V-D-J region (159-2B) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0580
R:Peeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0580
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <PEE>
A:Experimental source: day 19 fetal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 38.5%; Score 15; DB 2; Length 5;
Best Local Similarity 20.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 4 XXXWX 8
DB 1 ASSWD 5

RESULT 14
S65195
A:Title: alcohol dehydrogenase (EC 1.1.1.1) class III low affinity form - cod (Gadus sp.) (fragment)

C;Species: Gadus sp. (cod)
C;Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 12-Jun-1998
C;Accession: S66195
R;Hjeltnqvist, L.; Hackett, M.; Shafgat, J.; Danielsson, O.; Iida, J.; Hendrickson, R.C.;
FEBS Lett. 367, 237-240, 1995
A;Title: Multiplicity of N-terminal structures of medium-chain alcohol dehydrogenases. M
nzyme.

A;Reference number: S66191; MUID:95331382; PMID:7607314

A;Accession: S66195

A;Molecule type: protein

A;Residues: 1-6 <HJE>

C;Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

C;Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match 38.5%; Score 15; DB 2; Length 6;

Best Local Similarity 20.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 4 XXXWX 8

Db : : : :

2 AVAW 6

RESULT 15

PT0629

T-cell receptor beta chain V-D-J region (100-2AH) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C;Accession: PT0629; PT0528

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0629

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-6 <FEE>

A;Experimental source: newborn thymus, strain BALB/c, clone 100-2AH

A;Accession: PT0528

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-6 <FE2>

A;Experimental source: adult thymus, strain BALB/c, clone 100-4AB

C;Keywords: T-cell receptor

Query Match

38.5%; Score 15; DB 2; Length 6;

Best Local Similarity 50.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWX 4

Db : : : :

3 GDWG 6

Search completed: January 3, 2005, 17:36:16

Job time : 37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 17:26:09 ; Search time 188 Seconds
(without alignments)
24.484 Million cell updates/sec

Title: US-10-046-922-68
Perfect score: 39
Sequence: 1 GYWXWX 8

Scoring table: BLOSUM62DX, Gapop 10.0, Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 835

Minimum DB seq length: 0
Maximum DB seq length: 835

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	66.7	8	Q62721	rattus norv
2	26	66.7	8	AAB33374	homo sapi
3	25	64.1	5	UF01	MOUSE
4	18	46.2	8	Q64971	alfalfa mos
5	17	43.6	8	LK5	LEUMA
6	17	43.6	8	PK4	PERAM
7	17	43.6	8	Q7GEM6	branchiosto
8	17	43.6	8	Q8W8G2	diadema sav
9	17	43.6	8	Q8W8G3	diadema pau
10	17	43.6	8	Q8W8G4	diadema mex
11	17	43.6	8	Q8W8G5	diadema ant
12	17	43.6	8	Q8W8G6	diadema mex
13	16	41.0	4	OCF3	OCTMI
14	16	41.0	6	E101	LITRU
15	16	41.0	7	Q9BRV4	litoria rub
16	16	41.0	8	AK1	MELML
17	16	41.0	8	COM2	CONPU
18	16	41.0	8	LO2831	conus purpu
19	15	38.5	6	L0K1	LOCOMI
20	15	38.5	7	BRHP	CONIM
21	15	38.5	7	TPPY	PACDA
22	15	38.5	7	TY51	LITRU
23	15	38.5	7	Q95945	saccharomyc
24	15	38.5	7	Q49223	glycine max
25	15	38.5	7	Q8JE81	human immu
26	15	38.5	8	ACI	THUAL
27	15	38.5	8	AKG	GRYBI
28	15	38.5	8	AKH	LITAU
29	15	38.5	8	AKH	PROTE
30	15	38.5	8	AKH	TABAT
31	15	38.5	8	C125	CYPDO

32	15	38.5	8	1	CCKN	MACEU	P30369	macropus eu
33	15	38.5	8	1	HTF1	PERAM	P04548	periplaneta
34	15	38.5	8	1	HTF2	PERAM	P04549	periplaneta
35	15	38.5	8	1	HTF	TENMO	P25419	tenebrio mo
36	15	38.5	8	1	LCK1	LEUMA	P21140	leucophaea
37	15	38.5	8	1	LCK2	LEUMA	P21141	leucophaea
38	15	38.5	8	1	LCK3	LEUMA	P21142	leucophaea
39	15	38.5	8	1	LCK4	LEUMA	P21143	leucophaea
40	15	38.5	8	1	LCK6	LEUMA	P19988	leucophaea
41	15	38.5	8	1	LCK7	LEUMA	P19989	leucophaea
42	15	38.5	8	1	LCK8	LEUMA	P19990	leucophaea
43	15	38.5	8	1	PK1	PERAM	P82685	periplaneta
44	15	38.5	8	1	PK2	PERAM	P82686	periplaneta
45	15	38.5	8	1	PK3	PERAM	P82687	periplaneta

ALIGNMENTS

RESULT 1
ID Q62721 PRELIMINARY; PRT; 8 AA.
AC Q62721;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE Prohibition (Fragment)...
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fisher;
RX MEDLINE=95311633; PubMed=7607556;
RA Altus M.S., Wood C.M., Stewart D.A., Roskams A.I., Friedman V.,
RA Henderson T., Owens G.A., Danner D.B., Jupe E.R., Dell'Orco R.T.,
RA McClung J.K.;
RT "Regions of evolutionary conservation between the rat and human
prohibitin-encoding genes."
RL Gene 158:291-294(1995).
DR EMBL; U17178; AAA86692.1; -.
FT NON TER 8
SQ SEQUENCE 8 AA; 1150 MW; EFD3237B05A41376 CRC64;
Query Match 66.7%; Score 26; DB 2; Length 8;
Best Local Similarity 33.3%; Pred. NO. 1.8e+06;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 3 WXXXWX 8
Db |::|:
2 WRSEWK 7
RESULT 2
ID AAB33374 PRELIMINARY; PRT; 8 AA.
AC AAB33374;
DT 02-MAR-2004 (TREMBLrel. 27, Created)
DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)
DE 02-MAR-2004 (TREMBLrel. 27, Last annotation update)
DE Collagen alpha 5(IV) chain (Fragment).
GN COL4A5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95156893; PubMed=7853788;
RA Nakazato H., Hattori S., Ushijima T., Matsuura T., Koitabashi Y.,
RA Takada T., Yoshioka K., Endo F., Matsuuda I.;
RT "Mutations in the COL4A5 gene in Alport syndrome: a possible mutation

```

RT in primordial germ cells."
RL Kidney Int.. 46:1307-1314(1994).
DR EMBL; S75903; AAB33374.1; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 933 MW; 7370437735BAB378 CRC64;

Query Match 66.7%; Score 26; DB 2; Length 8;
Best Local Similarity 33.3%; Pred. No. 1.8e+06;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXWX 8
Db |:::|
2 WDSLWI 7

RESULT 3
UF01_MOUSE STANDARD; PRT; 5 AA.
ID UF01_MOUSE
AC P38639;
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Unknown protein from 2D-PAGE of fibroblasts (P19) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Fibroblast; PubMed=7523108;
RX MEDLINE=95009907;
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT "Separation and sequencing of familial and novel murine proteins using
RT preparative two-dimensional gel electrophoresis.";
RL Electrophoresis 15:735-745(1994).
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC protein is: 6.6, its MW is: 19 kDa.
KW Direct protein sequencing.
FT NON_TER 5
SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;

Query Match 64.1%; Score 25; DB 1; Length 5;
Best Local Similarity 40.0%; Pred. No. 1.8e+06;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXW 7
Db |:::|
1 WIGRW 5

RESULT 4
Q64971 PRELIMINARY; PRT; 8 AA.
ID Q64971
AC Q64971;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Putative ORF (Fragment).
OS Alfalfa mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Alfamovirus.
OX NCBI_TaxID=12321;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81124289; PubMed=6927843;
RA Koper-Zwarthoff E.C., Bredetode F.T.M., Veeneman G., van Boom J.H.,
RA Bol J.F.;
RT "Nucleotide sequences at the 5'-termini of the alfalfa mosaic virus
RT RNAs and the intercistronic junction in RNA 3.";
RL Nucleic Acids Res. 8:5635-5647(1980).
DR EMBL; V00047; CAA23416.1; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 917 MW; 69D40B0775A365B8 CRC64;

Query Match 46.2%; Score 18; DB 2; Length 8;
Best Local Similarity 33.3%; Pred. No. 1.8e+06;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWXW 6
Db |:::|
1 GWSWPE 6

RESULT 5
LCK5_LEUMA STANDARD; PRT; 8 AA.
ID LCK5_LEUMA
AC P19987;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Leucokinin V (L-V).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
OC Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE.
RC TISSUE=Head; PubMed=2877794;
RX MEDLINE=87052651; PubMed=2877794;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure, and synthesis of leucokinin V and VI:
RT myotropic peptides of Leucophaea maderae.";
RL Comp. Biochem. Physiol. 88C:27-30(1987).
CC -!- FUNCTION: This cephalomyotropic peptide stimulates contractile
CC activity of cockroach prothodum (hindgut).
CC -!- SUBCELLULAR LOCATION: Secreted.
DR PIR; JS0315; JS0315.
KW Amidation; Direct protein sequencing; Neuropeptide.
FT MOD_RES 8
FT SEQUENCE 8 AA; 784 MW; 736365A5B9C865B8 CRC64;

Query Match 43.6%; Score 17; DB 1; Length 8;
Best Local Similarity 25.0%; Pred. No. 1.8e+06;
Matches 2; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYWXW 8
Db |:::|
1 GSGFSSWG 8

RESULT 6
PK4_PERAM STANDARD; PRT; 8 AA.
ID PK4_PERAM
AC P82688;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Kinin-4 (Pea-K-4).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattoidea;
OC Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND AMIDATION.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=98010462; PubMed=9350979;
RA Predel R., Kellner R., Rapus J., Penzlin H., Gade G.;
RT "Isolation and structural elucidation of eight kinins from the
RT retrocerebral complex of the American cockroach, Periplaneta
RT americana.";
RL Regul. Pept. 71:199-205(1997).
CC -!- FUNCTION: Mediates visceral muscle contractile activity (myotropic
CC activity).
CC -!- SUBCELLULAR LOCATION: Secreted.

```



```

CC -!- MASS SPECTROMETRY: MW=838.15; METHOD=Electrospray; RANGE=1-8;
CC NOTE=Ref.1.
CC -!- SIMILARITY: Belongs to the kinin family.
CC Amidation; Direct protein sequencing; Neuropeptide.
FT MOD_RES 8 8 839 MW; 736365ASB9D6DD8 CRC64;
SQ SEQUENCE 8 AA; 839 MW; 736365ASB9D6DD8 CRC64;

Query Match 43.6%; Score 17; DB 1; Length 8;
Best Local Similarity 25.0%; Pred. No. 1.8e+06;
Matches 2; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYXXXXXX 8
| : : : :
Db 1 GAQFSSWG 8

RESULT 7
Q7GEM6 PRELIMINARY; PRT; 8 AA.
AC Q7GEM6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit I (Fragment).
GN Name=COI;
OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OC NCBI_TaxID=7740;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole organism;
RX MEDLINE=97398704; PubMed=9254918;
RA Delarbre C., Barriel V., Tillier S., Janvier P., Gachelin G.;
RT "The main features of the cranial mitochondrial DNA between the ND1
and the COI genes were established in the common ancestor with the
RT lancelet."
RL Mol. Biol. Evol. 14:807-813(1997).
DR EMBL; Y09524; CAA70710.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 8
SQ SEQUENCE 8 AA; 1129 MW; F0C7336411A04B56 CRC64;

Query Match 43.6%; Score 17; DB 2; Length 8;
Best Local Similarity 16.7%; Pred. No. 1.8e+06;
Matches 1; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXXX 8
| : : : :
Db 2 YITRWL 7

RESULT 8
Q8W8G2 PRELIMINARY; PRT; 8 AA.
AC Q8W8G2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Diadema savignyi (Longspine black urchin).
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diadematacea; Diadematoidea; Diadematiidae;
OC Diadema.
OC NCBI_TaxID=105360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;

us-10-046-922-68.dx-sz8.rup

```

```

RT "Population structure and speciation in tropical seas: global
RT phylogeography of the sea urchin Diadema.";
RL Evolution 55:955-975(2001).
DR EMBL; AY013065; AAL33860.1; -.
DR EMBL; AY013080; AAL33861.1; -.
DR EMBL; AY013083; AAL33862.1; -.
DR EMBL; AY013086; AAL33863.1; -.
DR EMBL; AY013088; AAL33864.1; -.
DR EMBL; AY013090; AAL33866.1; -.
DR EMBL; AY013091; AAL33867.1; -.
DR EMBL; AY013102; AAL33868.1; -.
DR EMBL; AY013103; AAL33869.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 8 AA; 1037 MW; 701B173B46DDC2D3 CRC64;

Query Match 43.6%; Score 17; DB 2; Length 8;
Best Local Similarity 16.7%; Pred. No. 1.8e+06;
Matches 1; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXXX 8
| : : : :
Db 1 WVAQYL 6

RESULT 9
Q8W8G3 PRELIMINARY; PRT; 8 AA.
AC Q8W8G3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Diadema paucispinum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diadematacea; Diadematoidea; Diadematiidae;
OC Diadema.
OC NCBI_TaxID=145530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;
RT "Population structure and speciation in tropical seas: global
RT phylogeography of the sea urchin Diadema.";
RL Evolution 55:955-975(2001).
DR EMBL; AY012960; AAL33853.1; -.
DR EMBL; AY012961; AAL33854.1; -.
DR EMBL; AY012962; AAL33855.1; -.
DR EMBL; AY012963; AAL33856.1; -.
DR EMBL; AY012964; AAL33857.1; -.
DR EMBL; AY012965; AAL33858.1; -.
DR EMBL; AY012966; AAL33859.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 8 AA; 1037 MW; 701B173B46DDC2D3 CRC64;

Query Match 43.6%; Score 17; DB 2; Length 8;
Best Local Similarity 16.7%; Pred. No. 1.8e+06;
Matches 1; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXXX 8
| : : : :
Db 1 WVAQYL 6

RESULT 10
Q8W8G4 PRELIMINARY; PRT; 8 AA.
AC Q8W8G4;

```

```

DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Diadema mexicanum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diadematacea; Diadematoidea; Diadematiidae;
OC Diadema.
OX NCBI_TaxID=105359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;
RT "Population structure and speciation in tropical seas: global
RT phylogeography of the sea urchin Diadema.";
RL Evolution 55:955-975(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21561594; PubMed=11703875;
RA Lessios H.A., Garrido M.J., Kessing B.D.;
RT "Demographic history of Diadema antillarum, a keystone herbivore on
RT Caribbean reefs.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 268:2347-2353(2001).
DR EMBL; AY012908; AAL33837.1; -.
DR EMBL; AY012911; AAL33838.1; -.
DR EMBL; AY012913; AAL33839.1; -.
DR EMBL; AY012914; AAL33840.1; -.
DR EMBL; AY012919; AAL33842.1; -.
DR EMBL; AY012940; AAL33847.1; -.
DR EMBL; AY012949; AAL33849.1; -.
DR EMBL; AY012950; AAL33850.1; -.
DR EMBL; AY012951; AAL33851.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER 1
SQ SEQUENCE 8 AA; 1037 MW; 701B173B46DDC2D3 CRC64;

Query Match 43.6%; Score 17; DB 2; Length 8;
Best Local Similarity 16.7%; Pred. No. 1.8e+06;
Matches 1; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXWX 8
DB |::: 1 WVAQYL 6

RESULT 11
Q8W8G5 PRELIMINARY; PRT; 8 AA.
AC Q8W8G5;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Diadema antillarum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diadematacea; Diadematoidea; Diadematiidae;
OC Diadema.
OX NCBI_TaxID=105358;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;
RT "Population structure and speciation in tropical seas: global
RT phylogeography of the sea urchin Diadema.";
RL Evolution 55:955-975(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21561594; PubMed=11703875;

Query Match 43.6%; Score 17; DB 2; Length 8;
Best Local Similarity 16.7%; Pred. No. 1.8e+06;
Matches 1; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXWX 8
DB |::: 1 WVAQYL 6

RESULT 12
Q8W8G6 PRELIMINARY; PRT; 8 AA.
AC Q8W8G6;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Diadema mexicanum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diadematoidea; Diadematiidae;
OC Diadema.
OX NCBI_TaxID=105359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;
RT "Population structure and speciation in tropical seas: global
RT phylogeography of the sea urchin Diadema.";
RL Evolution 55:955-975(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21561594; PubMed=11703875;
RA Lessios H.A., Garrido M.J., Kessing B.D.;
RT "Demographic history of Diadema antillarum, a keystone herbivore on
RT Caribbean reefs.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 268:2347-2353(2001).
DR EMBL; AY012931; AAL33845.1; -.
DR EMBL; AY012932; AAL33846.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER 1
SQ SEQUENCE 8 AA; 1076 MW; 701B173B46CAF2D3 CRC64;

Query Match 43.6%; Score 17; DB 2; Length 8;
Best Local Similarity 33.3%; Pred. No. 1.8e+06;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXWX 8
DB |::: 1 WVAQYL 6

RESULT 13
OCP3_OCTMI
ID_OCP3_OCTMI STANDARD; PRT; 4 AA.

```

```

RA Lessios H.A., Garrido M.J., Kessing B.D.;
RT "Demographic history of Diadema antillarum, a keystone herbivore on
RT Caribbean reefs.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 268:2347-2353(2001).
DR EMBL; AY012796; AAL33821.1; -.
DR EMBL; AY012799; AAL33822.1; -.
DR EMBL; AY012799; AAL33822.1; -.
DR EMBL; AY012803; AAL33823.1; -.
DR EMBL; AY012804; AAL33824.1; -.
DR EMBL; AY012806; AAL33825.1; -.
DR EMBL; AY012852; AAL33826.1; -.
DR EMBL; AY012854; AAL33828.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER 1
SQ SEQUENCE 8 AA; 1037 MW; 701B173B46DDC2D3 CRC64;

Query Match 43.6%; Score 17; DB 2; Length 8;
Best Local Similarity 16.7%; Pred. No. 1.8e+06;
Matches 1; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXWX 8
DB |::: 1 WVAQYL 6

RESULT 12
Q8W8G6 PRELIMINARY; PRT; 8 AA.
AC Q8W8G6;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Diadema mexicanum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diadematoidea; Diadematiidae;
OC Diadema.
OX NCBI_TaxID=105359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;
RT "Population structure and speciation in tropical seas: global
RT phylogeography of the sea urchin Diadema.";
RL Evolution 55:955-975(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21561594; PubMed=11703875;
RA Lessios H.A., Garrido M.J., Kessing B.D.;
RT "Demographic history of Diadema antillarum, a keystone herbivore on
RT Caribbean reefs.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 268:2347-2353(2001).
DR EMBL; AY012931; AAL33845.1; -.
DR EMBL; AY012932; AAL33846.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER 1
SQ SEQUENCE 8 AA; 1076 MW; 701B173B46CAF2D3 CRC64;

Query Match 43.6%; Score 17; DB 2; Length 8;
Best Local Similarity 33.3%; Pred. No. 1.8e+06;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 WXXXWX 8
DB |::: 1 WVAQYL 6

RESULT 13
OCP3_OCTMI
ID_OCP3_OCTMI STANDARD; PRT; 4 AA.

```

P58649;
 28-FEB-2003 (Rel. 41, Created)
 28-FEB-2003 (Rel. 41, Last sequence update)
 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Cardioactive peptides Ocp-3/Ocp-4.
 OS Octopus minor (Octopus).
 OC Eukaryota; Metazoa; Cephalopoda; Coleoidea; Neocoleoidea;
 OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
 OX NCBI_TaxID=89766;
 (1) NCBI_TaxID=89766;
 RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=20336815; PubMed=10876044;
 RA Iwakoshi E., Hisada M., Minakata H.;
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
 RT Octopus minor."; 630(2000).
 CC Peptides 21:623-630(2000).
 CC -!- FUNCTION: Cardioactive; has both positive chronotropic and
 CC inotropic effects on the heart. Ocp-4 is a 1000 time less active
 CC than Ocp-3.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PTM: Ocp-4 has D-Ser instead of L-Ser.
 CC -!- MASS SPECTROMETRY: MW=395.2; METHOD=WALDI; RANGE=1-4; NOTE=Ref.1.
 KW D-amino acid; Direct protein sequencing; Hormone.
 FT MOD_RES 2 2 D-serine (in form Ocp-4).
 SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;
 Query Match 41.0%; Score 16; DB 1; Length 4;
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GYWX 4
 DB 1 GSWD 4

RESULT 14
 E101 LITRU STANDARD; PRT; 6 AA.
 AC P82096;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Electrin 1.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylinae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=104895;
 (1) NCBI_TaxID=104895;
 RP SEQUENCE.
 RC TISSUE=Skin secretion;
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 RT Litoria rubella. Comparison with the skin peptides from Litoria
 RT rubella.";
 RL Aust. J. Chem. 52:639-645(1999).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Skin.
 KW Amidation; Amphibian defense peptide; Direct protein sequencing.
 FT MOD_RES 6 6 Methionine amide.
 SQ SEQUENCE 6 AA; 792 MW; 6683704772C9A000 CRC64;
 Query Match 41.0%; Score 16; DB 1; Length 6;
 Best Local Similarity 16.7%; Pred. No. 1.8e+06;
 Matches 1; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 QY 3 WXXXWX 8
 DB 1 FVPIMW 6

RESULT 15

Q9BRY4 PRELIMINARY; PRT; 7 AA.
 ID Q9BRY4;
 AC Q9BRY4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE SQSTM1 protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 (1) NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC005857; AA05857.3; -;
 DR GO; GO:0005829; C:cytosol; ISS.
 DR GO; GO:0019901; F:protein kinase binding; ISS.
 DR GO; GO:0042169; F:SH2 domain binding; ISS.
 DR GO; GO:0043130; F:ubiquitin binding; ISS.
 DR GO; GO:0016197; P:endosome transport; ISS.
 DR GO; GO:0007242; P:intracellular signaling cascade; ISS.
 DR GO; GO:0045944; P:positive regulation of transcription from P...; ISS.
 DR GO; GO:0008104; P:protein localization; ISS.
 DR GO; GO:0043122; P:regulation of I-kappaB kinase/NF-kappaB cas...; ISS.
 DR GO; GO:0006950; P:response to stress; ISS.
 DR InterPro; IPR000449; UBA.
 DR InterPro; IPR000433; Znf_ZZ.
 DR Pfam; PF00569; ZZ; 1.
 DR SMART; SM00165; UBA; 1.
 DR SMART; SM00291; Znf_ZZ; 1.
 DR PROSITE; PS50030; UBA; 1.
 DR PROSITE; PS01357; ZF_ZZ_1; UNKNOWN_1.
 DR PROSITE; PS50135; ZF_ZZ_2; 1.
 SQ SEQUENCE 7 AA; 775 MW; 737728769DDDD6F0 CRC64;
 Query Match 41.0%; Score 16; DB 2; Length 7;
 Best Local Similarity 66.7%; Pred. No. 1.8e+06;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GYW 3
 DB 5 GLW 7

Search completed: January 3, 2005, 17:35:33
 Job time : 189 secs

THIS PAGE BLANK (USPTO)